



UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA

JUAN PABLO PORTILLA LLERENA

LIGNINA EM ESPÉCIES DE SACCHARUM SPP.

**CAMINAS**

**2016**

**JUAN PABLO PORTILLA LLERENA**

**LIGNINA EM ESPÉCIES DE SACCHARUM SPP.**

*Dissertação apresentada ao Instituto de  
Biologia da Universidade Estadual de  
Campinas como parte dos requisitos  
exigidos para a obtenção do Título de  
Mestre em Biologia Vegetal*

ESTE ARQUIVO DIGITAL CORRESPONDE À  
VERSÃO FINAL DA DISSERTAÇÃO  
DEFENDIDA PELO ALUNO JUAN PABLO  
PORTILLA LLERENA E ORIENTADA PELO  
PROF. DR. PAULO MAZZAFERA

Orientador: PROF. DR. PAULO MAZZAFERA



**CAMPINAS**

**2016**

**Agência(s) de fomento e nº(s) de processo(s):** FAPESP, 2014/25994-1; CAPES

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca do Instituto de Biologia  
Mara Janaina de Oliveira - CRB 8/6972

Portilla Llerena, Juan Pablo, 1985-  
P836L Lignina em espécies de *Saccharum* spp. / Juan Pablo Portilla Llerena. –  
Campinas, SP : [s.n.], 2016.

Orientador: Paulo Mazzafera.  
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de  
Biologia.

1. *Saccharum*. 2. Carboidratos. 3. Lignina. 4. Parede celular. 5.  
Sacarificação. I. Mazzafera, Paulo, 1961-. II. Universidade Estadual de  
Campinas. Instituto de Biologia. III. Título.

Informações para Biblioteca Digital

**Título em outro idioma:** Lignin in species of the genus *Saccharum* spp.

**Palavras-chave em inglês:**

*Saccharum*

Carbohydrates

Lignin

Cell wall

Saccharification

**Área de concentração:** Biologia Vegetal

**Titulação:** Mestre em Biologia Vegetal

**Banca examinadora:**

Paulo Mazzafera [Orientador]

Camila Caldana

Michael dos Santos Brito

**Data de defesa:** 02-06-2016

**Programa de Pós-Graduação:** Biologia Vegetal

Campinas, 02 de Junho de 2016

### **COMISSÃO EXAMINADORA**

Prof. Dr. Paulo Mazzafera (orientador)

Profa. Dra. Camila Caldana

Prof. Dr. Michael dos Santos Brito

*Os membros da Comissão Examinadora acima assinaram a Ata de Defesa, que se encontra no processo de vida acadêmica do aluno.*

## DEDICATORIA

Dedico esta dissertação de mestrado...

A minha família...

*A minha querida mãe **Fabiola**, por estar comigo sempre nos momentos bons e maus. Por me ensinar a enfrentar com diligencia e sabedoria os desafios da vida e pelo apoio incondicional na minha formação acadêmica. Estou orgulhoso de você, por ser uma mulher com muita coragem e perseverança.*

*As minhas queridas irmãs: Karina, Carolina e Sofia, quanta saudade, carinho e afeto. Os momentos da vida são mais maravilhosos ao lado de vocês.*

***Jesuscristo*** é meu pastor e nada me faltará.  
Deitar-me faz em verdes pastos, guia-me  
mansamente a águas tranquilas. Refrigera a  
minha alma; guia-me pelas veredas da justiça,  
por amor do Seu nome.

Salmos 23:1-6

## **AGRADECIMENTOS**

Agradeço a Deus, Jesus Cristo e ao Espírito Santo, por ter me permitido caminhar até aqui, pela vida, por ter me dado força e paz, nos momentos difíceis.

Ao Prof. Dr. Paulo Mazzafera, pela recepção no seu laboratório e oportunidade de me orientar e me iniciar no longo caminho da ciência. Obrigado pela paciência e confiança desde o primeiro momento, pelas argumentações críticas e sugestões e por me ensinar que a ciência precisa sempre do rigor científico.

À Universidade Estadual de Campinas, particularmente ao Instituto de Biologia e ao Programa de Pós-Graduação em Biologia Vegetal, pela estrutura física para a realização desta pesquisa. A CAPES e a FAPESP com o processo 2014/25994-1 pela bolsa de mestrado concedida.

Ao Prof. Dr. Luis Ponce, pela amizade, confiança, pelo incentivo para estudar o mestrado em Brasil. O senhor é o grande responsável por eu estar aqui hoje. Sou grato pelos ensinamentos da vida e científicos transmitidos.

Aos membros da pre-banca e Banca, por terem aceitado o convite para participar da minha banca examinadora, e valorizado o trabalho realizado.

Aos Doutores: Eduardo Kiyota, Pedro Araújo e Luciano Pereira, pela ajuda nos experimentos de oligômeros de Lignina, análises moleculares e estadística, respetivamente...Muito obrigado amigos !!!

Ao Prof. Rafael Vasconcelos Ribeiro e a Ma. Fernanda Castro pela colaboração para a determinação de amido e a Prof. Dra. Juliana L. S. Mayer pela colaboração nas análises histoquímicas.

Ao Prof. Dr. Markus Pauly do Instituto de Biociências da Energia, Universidade de Berkeley, California-USA, pela colaboração com os análises da parede celular por espectroscopia 2D HSQC NMR.

Aos todos os colegas do Departamento de Fisiologia Vegetal, em especial aos membros do LaFiMP: Daniela, Ilse, Mariana, Vanessa, Nathalia, Rafaela, Flavia, Uiara, Leticia, Felipe, Caio, Luana, Dimitry, João e Tatiane que de alguma forma participaram na realização deste trabalho.

Aos novos amigos: Carlos e Ivan Rubianes e a Raul Lima pelas incontáveis horas de trabalho. Obrigado pelos bons momentos vividos.

Y de forma muito especial a Leia, Thais e Lana Santos de Souza e a Luiz Antônio de Oliveira Neto e Rodrigo Jaques de Souza, por ter encontrado neles sempre uma amizade sincera, fazendo com que meus dias fossem de muita alegria e tranquilidade, longe do meu país, Perú, e da minha família.

## RESUMO

Cana de açúcar é uma das plantas mais eficientes em coletar e converter a luz solar em energia química, apresentando alta capacidade de acumular biomassa. Recentemente, com o advento do desenvolvimento do etanol de segunda geração (E2G), ou etanol lignocelulósico, o bagaço tem recebido atenção como matéria prima para a produção desse biocombustível. O seu uso, porém, apresenta grandes desafios, principalmente porque a biomassa de cana é lignificada (~23%), situação que limita o processo de sacarificação. Portanto, o estudo dos fatores que governam a recalcitrância da parede celular em plantas que apresentam características de bioenergia é de relevante importância para o desenvolvimento de variedades híbridas de cana em programas de bioenergia, de tal forma que a produção de E2G possa ser maximizada. O objetivo desta pesquisa foi estudar a biossíntese de lignina em *Saccharum spontaneum*, *Saccharum officinarum*, *Saccharum robustum* e *Saccharum barberi*, através da identificação de genes da via biossintética da lignina e da análise dos componentes da parede celular em internódios novos (segundo + terceiro) e maduros (oitavo). Neste estudo, o conteúdo de carboidratos estruturais e não estruturais foi contrastante entre as espécies, uma vez que *S. officinarum* e *S. barberi* tiveram maior de acúmulo de açúcares solúveis totais, enquanto *S. spontaneum* e *S. robustum* apresentaram maior deposição de polissacarídeos estruturais. *S. spontaneum* apresentou um acúmulo notável de amido com relação às outras espécies de *Saccharum*. *S. spontaneum* e *S. robustum* apresentaram o maior conteúdo de glicose, enquanto xilose foi maior em *S. officinarum* e *S. barberi*. A determinação do conteúdo de lignina, em adição a testes histoquímicos com corantes específicos para lignina, confirmaram a maior deposição deste polímero nos internódios maduros das quatro espécies estudadas, sendo o processo de lignificação mais acentuado em *S. spontaneum* e *S. robustum*. A razão S/G e a porcentagem de sacarificação em internódios maduros foram menores em *S. spontaneum* e *S. robustum* do que em *S. officinarum* e *S. barberi*, e não apresentaram correlação, ou seja, a maior presença de unidades S não se refletiu em maior taxa de sacarificação. A análise do perfil de oligômeros por UPLC/MS foi possível identificar 11 estruturas, entre aldeídos, monômeros, dímeros e trímeros. Independentemente da idade do internódio, *S. spontaneum* e *S. robustum* possuem a maior diversidade e frequência de oligômeros quando comparadas com *S. officinarum* e *S. barberi*. A análise por espectroscopia 2D HSQC NMR evidenciou que *S. barberi* e *S. officinarum* possuem maior porcentagem de *p*-hidroxicinamatos (*p*-coumarato e ferulato) e grupos acetil substituintes (posição C-3 e total). Um total de 13 unigenes foram identificados nas espécies de *Saccharum*, os quais foram: 1 *COMT*, 1 *HCT*, 1 *F5H*, 2 *4CLs*, 3 *CAD*, 1 *C3H*, 1 *C4H*, 2 *CCoAOMT* e 1 *CCR*. A análise das árvores filogenéticas revelam que as sequências das enzimas da rota biossintética da lignina das espécies de *Saccharum* se agruparam preferencialmente, com o seus ortólogos de *Sorghum bicolor* e *Zea mays* e cana de açúcar (parentes mais próximos do gênero *Saccharum* na tribo Andropogoneae) em clados sempre suportados por alto valores de bootstrap. Os genes identificados apresentaram alta homologia entre si e com híbridos comerciais (Bottcher et al., 2013). Desta forma, os resultados mostram que existe conservação entre os genes entre espécies e híbridos, mas que de alguma forma sua regulação influencia o conteúdo e o tipo de lignina. Não foram feitas análises de expressão neste estudo, mas independente disso, os resultados mostram a importância de se estudar os fatores de transcrição no controle da biossíntese e de lignina, o que poderia explicar as variações encontradas no conteúdo e composição desse polímero da parede celular.



## ABSTRACT

Sugarcane ranks among the most efficient plants when it comes to collecting and converting sunlight into chemical energy, displaying a high capacity to accumulate biomass. With the advent of second generation ethanol (2GE), otherwise known as lignocellulosic ethanol, the bagasse has received a lot of attention as raw material for the production of said biofuel. Its use, however, poses great challenges, especially because the cane's biomass is lignified (~23%), which limits the saccharification process. Therefore, the study of the factors that govern the cell wall's recalcitrance in plants that present characteristics of bioenergy is relevant to the development of hybrid kinds of cane in bioenergetics programs in order to maximize the production of 2GE. The objective of this research has been to study biosynthesis of lignin in the *Saccharum spontaneum*, *Saccharum officinarum*, *Saccharum robustum* and *Saccharum barberi* species through the identification of genes of the biosynthetic pathway of lignin and the analysis of the components of the cell wall in both new (second + third) and mature (eighth) internodes. In this study, the content of both structural and non-structural carbohydrates has been contrasting between the different species, since both *S. officinarum* and *S. barberi* had a larger accumulation of total soluble sugars, while *S. spontaneum* and *S. robustum* had a larger deposition of structural polysaccharides. *S. spontaneum* presented a noticeable accumulation of starch in comparison with other species of *Saccharum*. *S. spontaneum* and *S. robustum* presented the largest glucose content, while the xylose was more abundant in *S. officinarum* and *S. barberi*. The determination of the lignin content, in addition to histochemical tests with lignin-specific dyes, has confirmed the larger deposition of this polymer in the mature internodes of the four studied species, the lignification process being more accentuated in *S. officinarum* e *S. barberi*. The S/G ratio and the saccharification percentage in mature internodes were smaller in *S. spontaneum* and *S. robustum* than in *S. officinarum* and *S. barberi*, and presented no correlation, that is, the bigger presence of S units did not translate into a higher saccharification ratio. From the analysis of the profile of the oligomers by UPLC/MS it was possible to identify 11 structures between aldehydes, monomers and trimers. Independently from the internodes' age, *S. spontaneum* and *S. robustum* possess the largest variety and frequency of oligomers when compared to *S. officinarum* and *S. barberi*. The analysis through 2D HSQC NMR spectroscopy has brought evidence that *S. officinarum* and *S. barberi* possess a larger percentage of *p*-hydroxycinnamates (*p*-coumarate and ferulate) and acetyl substituent groups (C-3 position and total). 13 unigenes in total have been identified in the *Saccharum* species, which were: 1 *COMT*, 1 *HCT*, 1 *F5H*, 2 *4CLs*, 3 *CAD*, 1 *C3H*, 1 *C4H*, 2 *CCoAOMT* and 1 *CCR*. The analysis of the phylogenetic trees revealed that the sequences of enzymes from the *Saccharum* species' lignin's biosynthetic pathway grouped preferentially with its orthologs of *Sorghum bicolor*, *Zea mays* and sugarcane (the most closely related to the *Saccharum* gender in the *Andropogoneae* tribe) into clades always supported by high bootstrap values. The identified genes presented high homology between themselves and with commercial hybrids (Bottcher et al., 2013). Thus, the results show that there is conservation of genes between species and hybrids, but their regulation influences somehow both the content and the kind of lignin. Expression analysis weren't made in this study, but the results show nevertheless the importance of studying the transcription factors in the control of lignin biosynthesis, which could explain the variations found in both content and composition of this polymer in the cell wall.

## Sumário

<b>CAMINAS 2016 .....</b>	<b>1</b>
<b>COMISSÃO EXAMINADORA .....</b>	<b>4</b>
<b>DEDICATORIA .....</b>	<b>5</b>
<b>AGRADECIMENTOS .....</b>	<b>7</b>
<b>RESUMO.....</b>	<b>8</b>
1.1. Etanol de primeira e segunda geração.....	11
1.2. Parede celular em gramíneas.....	12
1.3. Componentes da parede celular .....	13
1.4. Biossíntese da lignina .....	18
Figura 2. A provável história evolutiva das espécies de <i>Saccharum</i> (Grivet et al., 2006).....	24
1.6. Cana de açúcar e cana energia como fonte de biomassa lignocelulósica.....	25
<b>2. JUSTIFICATIVA .....</b>	<b>28</b>
<b>3. OBJETIVOS .....</b>	<b>29</b>
3.1. Objetivo Geral.....	29
3.2. Objetivos específicos.....	29
<b>4. MATERIAIS E MÉTODOS.....</b>	<b>30</b>
4.1. Material Biológico e condições de crescimento .....	30
4.2. Análises histoquímicas .....	31
4.3. Análises bioquímicas .....	31
4.4. Análises moleculares.....	37
4.5. Análises filogenéticas.....	39
4.6. Análises estatísticas.....	40
<b>5. RESULTADOS .....</b>	<b>41</b>
5.2. Carboidratos não estruturais .....	41
5.3. Lignina solúvel e insolúvel .....	43
5.4. Porcentagem de sacarificação.....	44
5.5. Razão S/G da lignina .....	44
5.6. Perfil de oligômeros da lignina .....	45
5.7. Fenóis totais .....	47
5.8. Composição de monossacarídeos, de lignina e acetilação da fração de xilano da parede celular.....	48
5.9. Análises histoquímicas - lignina, celulose e amido.....	50
5.10. Identificação de genes associados com a rota biossintética da lignina.....	55
5.10.1. Filogenia de C4H .....	56
5.10.2. Filogenia de 4CL .....	56
5.10.3. Filogenia de HCT.....	57
5.10.4. Filogenia de C3H .....	58
5.10.5. Filogenia de CCoAOMT .....	58
5.10.6. Filogenia de CCR.....	59
5.10.7. Filogenia de F5H.....	60
5.10.8. Filogenia de COMT .....	60
5.10.9. Filogenia de CAD .....	61
<b>6. DISCUSSÃO .....</b>	<b>63</b>
<b>7. CONCLUSÕES.....</b>	<b>75</b>
Figura S1 - C4H.....	76
Figura S1A-C4H .....	100
<b>5. REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>115</b>
Pinçon G, Chabannes M, Lapierre C, Pollet B, Ruel K, Joseleau J-P, Boudet AM,.....	123

## 1. INTRODUÇÃO

### 1.1. Etanol de primeira e segunda geração

O rápido crescimento populacional, bem como a industrialização global, exigem fornecimento contínuo e constante de energia (Li et al., 2008; Yuan et al., 2008). A redução das reservas de petróleo, o aquecimento global devido à emissão de gases de efeito estufa e outros problemas associados com o uso de combustíveis fósseis tem motivado a pesquisa dirigida ao desenvolvimento de fontes alternativas de energia, de preferência renovável, para a crescente demanda por energia em todo o mundo (Hisano et al., 2009; Waclawovsky et al., 2010). O principal biocombustível alternativo é o bioetanol, com uma produção anual que tem aumentado de 17 a  $86 \times 10^6 \text{ m}^3$ , no período de 2000 a 2011 (Ren, 2012), seguido do biodiesel, com uma produção anual de  $21.4 \times 10^6 \text{ m}^3$  em 2011.

O bioetanol pode ser produzido a partir do amido e da sacarose, denominado etanol de primeira geração (E1G), e biomassa lignocelulósica, chamado de etanol de segunda geração (E2G). O etanol do amido é produzido principalmente à partir do milho e, da sacarose, principalmente da cana de açúcar. Mais 10 bilhões de galões de E1G são produzidos no mundo a cada ano, através de processos industriais especializados, incluindo a hidrólise e fermentação (Goldemberg, 2007; Rass-Hansen et al., 2007). Embora a produção de E1G representa a opção mais conveniente e tecnicamente avançada, alguns tem sugerido que, a longo prazo, pode ocorrer uma competição entre as áreas destinadas ao plantio da cana de açúcar e culturas alimentícias (De Oliveira et al., 2005; Farrell et al., 2006). Para evitar esta situação no futuro, especialmente em locais que não dispõem de um clima favorável ou uma suficiente extensão territorial para cultivo, torna-se necessário investir no desenvolvimento de tecnologias para a produção de E2G. A tecnologia do E2G envolve o uso da biomassa vegetal, que pode servir como matéria-prima renovável para a produção de energia, podendo ser transformada em biocombustíveis através de um processo chamado sacarificação. Assim, a biomassa remanescente de uma cultura, a biomassa lignocelulósica, como por exemplo o bagaço de cana, seria uma grande fonte de açúcares para a produção etanol (Carroll & Somerville, 2009; Simmons *et al.*, 2008; Somerville *et al.*, 2010). Atualmente, apenas insignificantes quantidades de E2G são produzidos em unidades industriais ao redor do mundo, indicando que ainda esta tecnologia não está comercialmente factível, sendo que as limitações são tanto econômicas e técnicas, estas últimas decorrentes de problemas na otimização do processo (Lennartsson et al., 2014). O principal desafio técnico está

relacionado à recalcitrância da lignina, isto é, os polissacarídeos da parede celular não são prontamente disponíveis ao processo de sacarificação porque a lignina liga-se à celulose e à hemicelulose, protegendo-as fisicamente (Anderson and Akin, 2008; Buckeridge et al., 2008). Esta situação faz com o acesso à celulose por enzimas hidrolíticas seja dificultado assim como o tempo de degradação da biomassa seja longo, elevando o custo do processo (Soccol et al., 2010). Para resolver este problema é necessário submeter à matéria prima a tratamentos químicos prévios, removendo a lignina, reduzindo a cristalinidade da celulose e aumentando a porosidade do material, facilitando assim o processo de sacarificação. Embora os pré-tratamentos químicos existentes sejam relativamente eficientes para decompor a parede celular, existe a consequente geração de resíduos químicos que, além de serem potencialmente poluentes, produzem produtos que inibem a fermentação dos açúcares liberados, tais como os furfurais (Buckeridge et al., 2010). Alternativamente, para superar o problema do pré-tratamento com produtos químicos, tem sido proposto o desenvolvimento de plantas com menor recalcitrância, seja por melhoramento genético ou pela produção de plantas geneticamente modificadas, com alteração no conteúdo ou a composição de lignina através da super-expressão ou silenciamento de genes que codificam enzimas que participam na biossíntese deste polímero (Ralph et al., 2006; Chen and Dixon, 2007).

## **1.2. Parede celular em gramíneas**

A parede celular é uma estrutura complexa e dinâmica, presente em quase todas as células vegetais. Esta estrutura está relacionada a processos fisiológicos, como dar forma e tamanho às células, conferir resistência mecânica aos tecidos, controlar a expansão celular, atuar no transporte intercelular, participar da sinalização e do reconhecimento entre células, armazenar compostos de reserva, e também moléculas reguladoras e sinalizadoras que controlam diversos processos fisiológicos celulares, além de participar dos mecanismos de proteção contra os microrganismos (Buckeridge et al., 2008; Sousa, 2011). Dois tipos de parede celular são encontrados em fanerógamas, paredes primárias e paredes secundárias. As primárias são produzidas por células em crescimento e, portanto, podem se alongar; as paredes secundárias não tem essa capacidade por terem sua elasticidade reduzida devido ao acúmulo de todos seus componentes, em especial lignina, que não se acumula na parede primária (Hayashi, 1989). Tipicamente, as paredes celulares são compostas pelos polissacarídeos celulose, hemicelulose e pectina (Scheller and Ulvskov, 2010a) e por lignina, um polímero de monolignóis derivados da via dos fenilpropanóides (Liepman et al., 2010).

Nos modelos mais recentes é proposto que os polímeros da parede celular formam três domínios independentes: celulose-hemicelulose, pectinas e proteínas (McCann and Roberts, 1991; Carpita and Gibeaut, 1993). O domínio celulose-hemicelulose influencia na forma celular enquanto o domínio péctico determina o grau de porosidade da parede, fornece moléculas sinalizadoras para disparar o mecanismo de defesa e armazena cálcio ao longo da planta. O domínio protéico por sua vez, tem papel estrutural e enzimático. Dependendo das proporções entre estas classes e dos componentes das hemiceluloses e pectinas, as paredes podem ser classificadas em tipo I, II e III (Carpita and Gibeaut, 1993). A parede do tipo I, característica de eudicotiledôneas e monocotiledôneas não-comelinóides, é definida como rica em xiloglucano, como o principal componente da hemicelulose. Recentemente foi descoberto em samambaias a parede do tipo III, com hemicelulose composta principalmente por mananos e com baixa quantidade de pectinas (Silva et al., 2011). Em contraste às paredes do tipo I e III, a parede de cana de açúcar, assim como de outras gramíneas, é considerada do tipo II, cujos componentes principais da hemicelulose são o  $\beta$ -glucano (BG) e arabinoxilano (AX), e tem o domínio péctico reduzido. O xiloglucano (XG) é encontrado em baixas concentrações. Além disso uma quantidade relativamente alta de composto fenólicos (17-24%) é encontrada nas paredes celulares de tipo III, interligando os polissacarídeos, mas não necessariamente na forma de lignina (da Silva, 2005).

### 1.3. Componentes da parede celular

**Celulose:** O principal polissacarídeo da parede celular vegetal é a celulose, um homopolímero linear formado por resíduos de D-glucopiranoses interligados por ligações do tipo  $\beta$ 1-4, com grupamentos hidroxilas nos carbonos 2, 3 e 6 (Klemm et al., 1998). A solubilidade do polímero de anhidroglicose em água diminui acima do grau de polimerização 6, ou seja, 6 resíduos, devido às fortes ligações de hidrogênio intermoleculares. A celulose de bagaço da cana de açúcar tem um massa molecular entre 150.000 e 350.000 Da (Pena et al., 2000), os quais equivalem a um grau de polimerização entre 800 e 1.900. No nível supramolecular, a composição química e a conformação espacial da molécula de celulose resulta numa tendência à agregação, com estruturas altamente ordenadas, através de uma extensiva rede de pontes de hidrogênio. Esta agregação estrutural não é uniforme, apresentando regiões de alta e de baixa cristalinidade, ou estado amorfo (Klemm et al., 1998). A interação de 36 moléculas desse polissacarídeo através de pontes de hidrogênio forma aglomerados moleculares com

estrutura para-cristalina denominado microfibrillas. Acredita-se que a orientação destas microfibrillas tenha um papel importante na manutenção do formato da célula vegetal e também no alongamento celular. A orientação das microfibrillas é coordenada pelas hemiceluloses, cuja composição e estrutura variam conforme o grupo taxonômico ao qual pertence a planta (Buckeridge et al., 2008).

**Hemicelulose:** As hemiceluloses são polímeros heterogêneos de pentoses (xilose e arabinose), hexoses (manose, glicose e galactose) e ácidos urônicos (Saha, 2003). As hemiceluloses são tipicamente ramificadas, com menor grau de polimerização do que a celulose, sendo entre 80 e 200 (Sun et al., 2004). As hemiceluloses se apresentam na forma de cadeias ramificadas (Ramos, 2003), não apresentam cristalinidade e como resultado são higroscópicas e mais acessíveis a hidrólise do que a celulose (Spiridon and Popa, 2005). Nas paredes celulares, as moléculas de hemicelulose estão ligadas por pontes de ferulato às microfibrillas de celulose (Laureano-Perez et al., 2005).

Na parede celular de cana de açúcar as hemicelulose são heteropolímeros, tendo na cadeia principal predominantemente D-xilose, com grupos laterais contendo principalmente ácido glucorônico e arabinose. A massa molecular da hemicelulose da cana de açúcar está entre 10.000 e 20.000 Da (Pena et al., 2000). Existe uma considerável variação nas proporções dos constituintes da hemicelulose, com uma razão molar de xilose para arabinosa de 4.0 a 52.6 e de xilose para ácido glucorônico de 7.4 a 100 (Lavarack et al., 2002). A extração da hemicelulose a partir do bagasso de cana com água a temperaturas entre 150°C e 170°C resulta numa produção de 60% de xilose sendo que 80% da xilose extraída está na forma de oligo- ou polissacarídeo (Saska and Ozer, 1995).

**Pectinas:** As pectinas possuem dois constituintes fundamentais: os ácidos poligalacturônicos (PGA), que são homopolímeros do ácido (1→4)-α-D-galactosilurônico (GalA), e ramnogalactarunanos do tipo I (RG I), os quais são heteropolímeros de repetidas unidades do dissacarídeo (1→2) α-L-ramnosil-(1→4)-α-D-GalA (Jarvis, 1984). RG I se encontra em paredes de células somáticas de milho e arroz (Thomas et al., 1989). Por outro lado, PGA e RG I se encontram em gramíneas, mas em pequenas quantidades (Shibuya and Nakane, 1984; Carpita, 1989). Assim como em muitas angiospermas, as gramíneas contêm arabinanos, galactanos e os altamente ramificados arabinogalactanos (Ags) de várias configurações e tamanhos e eles estão ligados à posição 0-4 dos resíduos ramnosil de RG (Shibuya and Nakane, 1984; Carpita, 1989). Os arabinanos são estruturas principalmente unidas por o 5-

arabinofuranosil, mas podem se unir os uns aos outros em praticamente cada posição livre (*O*-2, *O*-3 e *O*-5) e formar um diverso grupo de arabinanos ramificados. Os galactanos e duas classes de Ags são os principais grupos laterais. Uma classe de Ags está associada somente com as pectinas e é composto por fitas de (1→4)-β-D-galactanos, principalmente com t-arabinosil e algumas vezes com unidades de t- galactosil, anexados ao *O*-3 dos resíduos de galactosil da estrutura central do polímero (Bacic et al., 1988). Em pectinas de milho e arroz, os ramificados 5-arabinanos e as fitas laterais de 4-galactano estão unidas ao *O*-4 de cada dois de três unidades de ramnosil (Jarvis, 1984; Shibuya and Nakane, 1984; Thomas et al., 1989).

A segunda classe de Ags constitui um amplo grupo de cadeias curtas de (1→3)- e (1→6)-β-D-galactanos unidas umas a outras por resíduos (1→3,1→6) ramificados. A maioria das posições restantes, *O*-3 ou *O*-6 do galactosil, se encontram unidas aos grupos t-arabinosil (Fincher et al., 1983; Bacic et al., 1987). Esses Ags são também associados com proteínas (AGPs), que provavelmente tem funções intracelulares, na membrana plasmática e na parede celular, mas ainda não totalmente conhecidos.

**Lignina:** A lignificação das paredes celulares vegetais evoluiu há 400 milhões de anos, como uma das primeiras mudanças necessárias das plantas superiores para poder colonizar os diversos ambientes da superfície terrestre. A lignina é um polímero heterogêneo com propriedades únicas (Ralph et al., 2004), que se apresenta predominantemente no xilema da maioria das plantas terrestres, é o segundo biopolímero com maior abundância sobre a terra apresentando pelo menos o 30% de carbono da biomassa vegetal (Pauly and Keegstra, 2008). A via biosintética desta substância é complexa, e em termos de energia, a lignificação é um dos produtos mais caros feitos pelas plantas.

A lignina é uma molécula fenólica produzida por processos enzimáticos e por acoplamento de radicais livres de precursores monolignóis, processos que ocorrem na parede celular, depois de que o crescimento vegetal tem terminado e continua através da formação da parede celular secundária. O acoplamento radical resulta numa diversidade de ligações, e como a molécula de lignina cresce dentro de toda a parede, se pensa que a lignina se intercala entre os polissacarídeos. O caráter hidrofóbico da lignina mantém a água fora da parede celular, conferindo propriedades físicas e químicas únicas. A lignina é principalmente sintetizada em tecidos envolvidos no transporte de água ou no suporte mecânico, incluindo elementos de vaso no xilema, traqueídes, fibras, fibras floemáticas, esclerênquima e periderme.

A lignificação confere força e suporte mecânico, o que confere resistência à pressão negativa desenvolvida no xilema devido à transpiração (Jones et al., 2001), e também a resistência a insetos e patógenos (Franke et al., 2002). A lignificação também dificulta a digestibilidade da parede celular por herbívoros (Moore and Jung, 2001) e pode atuar como uma união intermolecular e intramolecular, interpenetrando e fortalecendo os carboidratos estruturais. As plantas com reduzidos conteúdos de lignina fracassam na capacidade para manter o corpo vegetal erguido ou se defender dos patógenos (Franke et al., 2002).

A lignina está covalentemente unida a uma variedade de polissacarídeos da parede (arabinoxilanos, celulose e pectina), por interações éster entre os grupos hidroxilas dos monolignóis e polissacarídeos, e por interações éter entre os grupos hidroxilas dos monolignóis e polissacarídeos ou pontes de ácido dihidroferúlico (Iiyama et al., 1994).

Os fenilpropanóides usados como blocos de construção para a molécula da lignina, são chamados monolignóis e são de três tipos: álcool *p*-coumaril, álcool coniferil e álcool sinapil, e diferem simplesmente no número de grupos metoxi ligados ao anel aromático, sendo os grupos metoxi os possíveis ponto de ramificação da lignina. Após o acoplamento oxidativo, esse álcoois dão, respectivamente, origem às subunidades: *p*-hidroxifenil (H), guaiacil (G) e siringil (S) (Ralph et al., 2004). Mais recentemente foi descrito que tricine é um importante componente da lignina, estando presente na parede celular de várias plantas (Lan et al., 2015). Embora sejam estes poucos monômeros os constituintes da lignina, a sua composição e estrutura varia grandemente entre as espécies.

A lignina é usualmente descrita pela relativa abundância e a razão entre suas unidades, particularmente G e S (Campbell and Sederoff, 1996). Com algumas notáveis exceções, as ligninas de gimnospermas são compostas somente de unidades G (com menor quantidades de unidades H), enquanto ligninas de gramíneas, incluindo cana de açúcar (He and Terashima, 1991), comprometem maiormente unidades S e G, com menos abundância de unidades H. As ligninas de muitas culturas podem ser aciladas com diferentes ácidos orgânicos (Ralph et al., 2004), mas a bioquímica e a função da acilação das ligninas, até o momento não tem sido elucidado. Os monolignóis são acilados para produzir éster-conjugados anteriormente à incorporação dentro de lignina (Lu and Ralph, 2002).

As unidades formadoras da lignina podem ser unidas por 20 tipos de ligações (Sun et al., 2000) que podem ser do tipo éter, envolvendo oxigênio ( $\beta$ -O-4',  $\alpha$ -O-4, 4-O-5') ou por meio de ligações condensadas envolvendo C-C ( $\beta$ - $\beta$ ',  $\beta$ -5', 5-5' e  $\beta$ -1'). Além disto, a mesma pode estar ainda associada com hemicelulose, através de ligações benzil-éter, benzil-



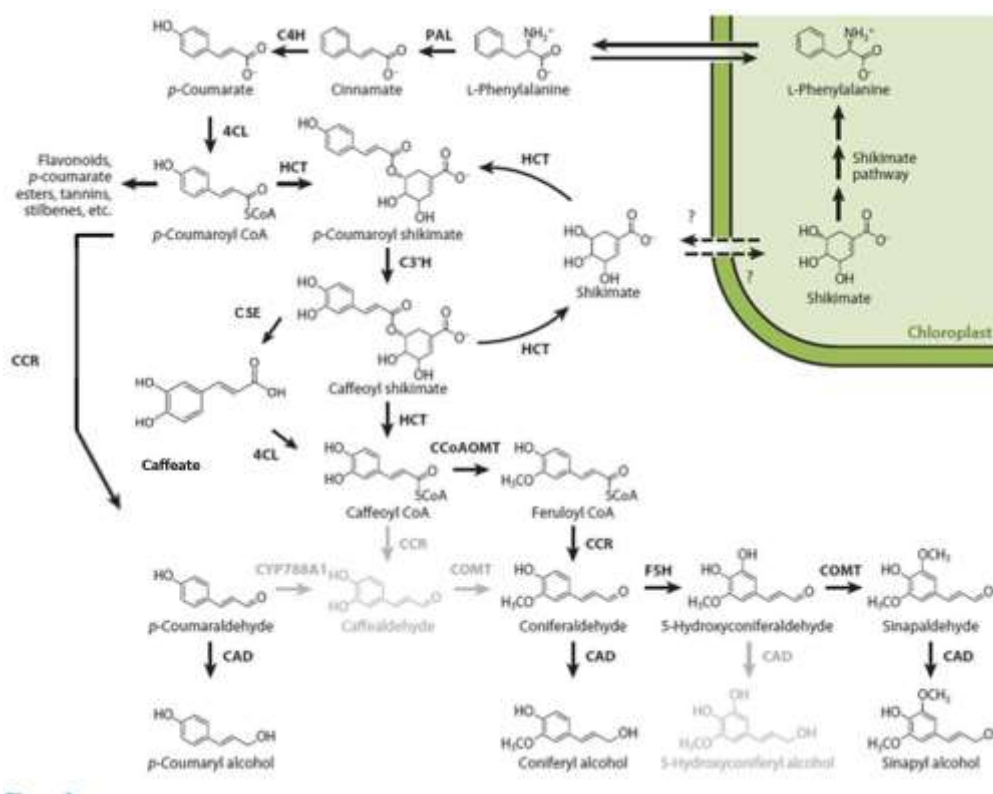
éster e fenil-glicosídeos (Guerra et al., 2006; Xu et al., 2008). As unidades constituídas por ligações  $\beta$ -O-4, denominadas  $\beta$ -aril éteres, são subestruturas bem frequentes na molécula de lignina (Sun et al., 2012), sendo predominantes nas ligninas de *Arabidopsis thaliana* (L.) Heynh., as quais também possuem 5 grandes quantidades de resinóis ( $\beta$ - $\beta$ ; Marita et al. 1999). Outras ligações bem comuns nas ligninas constituem as fenilcumaranas ( $\beta$ -5'; Rencoret et al. 2009). A lignina de gramíneas é unicamente acilada com *p*-coumarato no carbono terminal (C9) das subunidades de lignina (Lu and Ralph, 1999). Em milho o *p*-coumarato está preferencialmente acilado a unidades S (Grabber et al., 1996). Além das diferenças na composição de lignina entre espécies, a sua composição pode também diferir entre tecidos e também varia em função de vários fatores bióticos e abióticos (Moura et al., 2010). Em cana tem sido observado que estresse por falta de água (Santos et al., 2015) ou por temperatura (dos Santos et al., 2015) afeta não só a quantidade e composição da lignina, mas como também a expressão gênica dos genes relacionados à sua biossíntese.

**Substâncias aromáticas:** Uma principal característica das Poales e os seus clados mais próximos é o enriquecimento com substâncias aromáticas nas paredes celulares que ainda não estão lignificadas (Harris and Hartley, 1980; Rudall and Caddick, 1994). Uma grande porção das substâncias aromáticas são os ésteres dos hidroxicinamatos, ferulato e *p*-coumaratos (Harris and Hartley, 1980). Os glucoronoarabinoxilanos (GAX) estão entre-cruzados nas paredes celulares, tanto por hidroxicinamatos esterificados e eterificados e por outras substâncias fenólicas (Scalbert et al., 1985; Iiyama et al., 1993), e as uniões eterificadas representam complexos de polissacarídeos e lignina (Iiyama et al., 1990). Os ésteres de ferulato e *p*-coumarato estão unidos ao O-5 dos resíduos de arabinosil do GAX (Gubler and Ashford, 1985). (Markwalder and Neukom, 1976) sugerem que fitas vizinhas de GAX feruloizadas estão entrecruzadas pela formação de 5,5-diferulato, sendo isto demonstrado diretamente no GAX de bamboo (Ishii, 1991). O 5,5-diferulato é um dehidrodímero de toda uma série presente em gramíneas (Ralph et al., 1994b). Esses ésteres são quebrados por ataque com álcali diluído, e a liberação de ácido ferúlico e diferúlico está em concordância com a liberação de HS-GAX (Carpita, 1986). Os polissacarídeos podem também ser entrecruzados fotoquimicamente por [2 + 2]-homo-e-heterodimerização dos ésteres de ferulato e *p*-coumarato para derivados de ciclobutano chamados ácidos truxílicos e truxínico (Thomas et al., 1989; Turner et al., 1993).

Mais interações complexas entre substâncias aromáticas e polissacarídeos envolvem interações éster-éter e as ligações éter não são quebradas por álcali diluído (Iiyama et al., 1994). Sugere-se que tais interações éster-éter formam pontes entre polissacarídeos e lignina via adição de fenol para intermediários de lignina-quinona de vias diferentes. O Ácido ferúlico é o principal componente na ligação éster-éter de carboidratos e lignina (Iiyama et al., 1990). Os ferulatos também estão incorporados dentro de ligninas via mecanismos de formação de radicais, formando não somente as ligações  $\beta$ -eter, se não também estruturas que não podem ser liberadas por qualquer método solvolítico (Ralph et al., 1992). O ácido *p*-coumarico e ferúlico formam ésteres simples, mas raramente o ácido *p*-coumarico forma ligações éster-éter (Lam et al., 1992). O ácido *p*-coumarico é mais fortemente associado a lignina, particularmente nas últimas fases do desenvolvimento da parede celular e está unido às  $\gamma$ -posições das fitas laterais da lignina (Ralph et al., 1994).

#### **1.4. Biossíntese da lignina**

A biossíntese dos monolignóis começa com a deaminação da fenilalanina e envolve sucessivas reações de hidroxilação do anel aromático, seguidas por *O*-metilação fenólica e conversão das fitas carboxílicas laterais para um grupo álcool. A biossíntese dos monolignóis pode ser entendida a partir da rota metabólica apresentada na Figura 1. Pesquisadores tem acreditado por muito tempo que as reações de hidroxilação e a metilação ocorrem ao nível dos ácidos cinâmicos e que o ácido *p*-coumarico, ferúlico e sinápico são subsequentemente convertidos para os correspondentes monolignóis pela ação sequencial da enzima 4-coumarato:CoA ligase (4CL), cinamoil-CoA redutase (CCR) e cinamil álcool desidrogenase (CAD).



**Figura 1.** Rota simplificada da biossíntese da lignina, adaptada de Bonawitz & Chapple (2010). As enzimas envolvidas na via são: fenilalanina amônia-liase (PAL), cinamato 4-hidroxilase (C4H), 4-hidroxicinamoil CoA ligase (4CL), hidroxicinamoil CoA: shiquimato/quinato hidroxicinamoiltransferase (HCT), cafeoil shiquimato esterase (CSE), *p*-Coumarato 3-hidroxilase (C3H), cafeoil-CoA *O*-metiltransferase (CCoAOMT), cinamoil-CoA redutase (CCR), ferulato 5-hidroxilase (F5H), cafeato *O*-metiltransferase (COMT), cinamil álcool desidrogenase (CAD).

A primeira revisão da rota biossintética surge da observação que a expressão de cafeoil-CoA *O*-metiltransferase (CCoAOMT), coincide com a deposição da lignina em elementos traqueais em *Zinnia* (Ye et al., 1994). CCoAOMT metila cafeoil-CoA e 5-hidroxi feruloyl-CoA para produzir feruloyl-CoA e sinapoyl-CoA *in vitro*. Pesquisas futuras sugerem a existência de uma rota de *p*-coumaroyl-CoA para cafeoil-CoA (Zhong et al., 1998). Subsequentemente, experimentos com glucósidos de monolignóis marcados radioativamente amostram que a hidroxilação e metilação das posições aromáticas C<sub>3</sub> e C<sub>5</sub> poderiam ocorrer ao nível de aldeído e álcool (Chen et al., 1999), indicando a existência de enzimas capazes de catalisar essas conversões. Reavaliando a preferência do substrato da ferulato 5-hidroxilase (F5H), se mostra que o substrato preferencial não é o ácido ferúlico, mas sim o coniferaldeído e o álcool coniferil (Humphreys et al., 1999). Similarmente, os

produtos de hidroxilação, catalisados por F5H: 5-hidroxiciniferaldeído e 5-hidroxiciniferil álcool provam ser bons substratos para COMT, enquanto ácido caféico foi um pobre substrato (Humphreys et al., 1999). Juntamente, esses dados apresentam evidência que a posição C<sub>5</sub> do anel aromático está hidroxilada e metilada preferencialmente ao nível de cinamaldeído ou álcool cinamil e que o predominante papel da CCR é a redução do feruloil-CoA para ciniferaldeído. Numa variedade de espécies, cafeil aldeído e cafeil álcool são também eficientemente *O*-metilados, de acordo com estudos com traçadores radioativos (Chen et al., 2001). Cafeilaldeído é provavelmente sintetizado desde cafeoil-CoA por CCR (Guo et al., 2002).

Coniferaldeído pode ser sintetizado a partir do feruloil-CoA ou cafeilaldeído, dependendo das diferenças na especificidade de substrato entre as diferentes isoformas e as espécies (Zubieta et al., 2002). CAD é uma enzima multifuncional que catalisa a redução final dos cinamaldeídos para os correspondentes álcoois, pelos menos *in vitro*. Porém, uma CAD homóloga de *Populus*, sinapil álcool desidrogenase (SAD), que preferencialmente reduz sinapaldeído para álcool sinapil foi identificada. Este homólogo é co-expresso com F5H e COMT e contribui com ligninas S em *Populus* (Li et al., 2001). CAD de *Populus* preferencialmente reduz coniferaldeído, além disso, SAD pode ser a enzima responsável para o passo final na biossíntese do álcool sinapil (Li et al., 2001). A enzima que converte o ácido *p*-coumarico em ácido caféico foi clonada de *Arabidopsis* e mostrou ser um citocromo P450-monooxigenase (Schoch et al., 2001; Franke et al., 2002). É interessante notar que este ensaio enzimático tem demonstrado que os ésteres de shiquimato e quinato do ácido *p*-coumarico são os substratos preferencialmente para *p*-coumarato 3-hidroxilase (C3H) e que nem o ácido *p*-coumarico, *p*-coumaroil-CoA, *p*-coumaraldeído, *p*-coumaril álcool e nem o éster de 1-*O*-glicose e o 4-*O*-glicosido do ácido *p*-coumarico são bons substratos (Franke et al., 2002). Pela incorporação da C3H no esquema da biossíntese dos monolignóis, pelo menos em *Arabidopsis*, *p*-coumarato é primeiramente convertido a *p*-coumaroil-CoA pela 4CL, com a subsequente conversão para *p*-coumaroil-shiquimato e *p*-coumaroil quinato, os substratos da C3H, por *p*-hidroxicinamoil-CoA:D-quinato (CQT) ou *p*-hidroxicinamoil-CoA:shiquimato *p*-hidroxicinamoiltransferase (CST) (Schoch et al., 2001). Essas enzimas, descritas como reversíveis, podem converter cafeoil-shiquimato ou cafeoil-quinato (ácido clorogênico) em cafeoil-CoA, o substrato da CCoAOMT (Ulbrich and Zenk, 1980). Uma aciltransferase reversível com atividade tanto de CQT como CST, designada HCT, foi purificada de tabaco e o gene correspondente foi clonado (Hoffmann et al., 2003). Tomando em conjunto esses

dados, se argumenta que nenhum das substituições no C<sub>3</sub> e C<sub>5</sub> do anel aromático tem lugar ao nível do ácido cinâmico na biossíntese dos monolignóis.

Para a maioria das enzimas descritas acima, existem múltiplas isoformas que são diferencialmente expressas durante o desenvolvimento e sob diferentes condições ambientais (Chen et al., 2000), e podem ter diferentes cinéticas e preferência por diversos substratos (Ehlting et al., 1999). Certamente os diversos caminhos presentes na rota metabólica são esperados a ser cineticamente favorecidos por um dado tipo celular ou condições ambientais, permitindo uma plasticidade e flexibilidade metabólica. Neste sentido, uma continua controvérsia concerne o papel do ácido sinápico na biossíntese dos monolignóis. Considerando que as isoformas da 4CL das maioria das espécies vegetais analisadas usam *p*-coumarato, cafeato e ferulato como substratos, mas não sinapato, isoformas de algumas plantas são capazes de converter sinapato dentro de sinapoil-CoA (Lindermayr et al., 2002), deixando a possibilidade que em certas plantas, os monolignóis pode ser sintetizado pela via dos ácidos (Yamauchi et al., 2002). Outro nível de complexidade é que os intermediários da via podem afetar a síntese ou atividade de certas enzimas. Ácidos cinâmicos inibem a expressão da PAL (fenilalanina amônia-liase) ao nível transcricional (Blount et al., 2000) e ao nível pós-traducional (Bolwell et al., 1986) induz a atividade da CQT (Lamb, 1977). De acordo com a supressão da cinamato-4-hidroxilase (C4H) em plantas transgênicas de tabaco, se vê reduzida a atividade de PAL por modulação da retroalimentação (Blount et al., 2000). Concentrações de fenilalanina também tem um profundo efeito sobre o fluxo na rota biossintética. Incrementando a lignificação de células em suspensão de *Pinus taeda* com fenilalanina, houve aumento da síntese de *p*-coumaril, coniferil álcool e níveis de transcritos de PAL, 4CL, CCoAOMT e CCR, mas só fracamente C3H e C4H (Anterola et al., 2002). Experimentos *in vitro* tem mostrado que 5-hidroxiconiferaldeído é um inibidor competitivo da metilação do ácido cafeico e ácido 5-hidroxiferúlico e que coniferaldeído é um inibidor não competitivo da hidroxilação do 5-ferulato, corroborando a conclusão que a substituição nas posições C<sub>3</sub> e C<sub>5</sub> do anel aromático na biossíntese dos monolignóis não tomam lugar ao nível do ácido cinâmico (Li et al., 2000). Em *Populus*, coniferaldeído é adicionalmente um inibidor competitivo da redução do sinapaldeído por CAD, enquanto sinapaldeído inibe fortemente a redução do coniferaldeído por SAD, indicando que SAD é predominantemente envolvida na biossíntese de sinapil e CAD em coniferil álcool (Li et al., 2001). Ácido caféico inibe competitivamente a ativação de 4CL pelo ácido *p*-coumarico (Harding et al., 2002) e uma regulação por retroalimentação por níveis de ácido hidroxicinâmicos também tem sido

sugerida para CCoAOMT (Jin et al., 2000). Além disso, plantas transgênicas tem mostrado que a supressão de COMT em tabaco induz fortemente a atividade de CCR e tem um efeito negativo sobre a produção de CCoAOMT (Pinçon et al., 2001), indicando adicionais mecanismos regulatórios controlando o fluxo a través da rota biossintética.

Os três *p*-hidroxicinamil álcoois resultantes da rota biossintética dos monolignóis são transportados como conjugados desde o citosol até o apoplasto, onde eles experimentar uma polimerização desidrogenativa principalmente por peroxidases e/ou lacases para formar as ligninas (Boerjan et al., 2003). O modo da condensação dos monolignóis é controverso, por que tem sido considerado que isso ocorre por um processo de acoplamento por radicais, mas também tem sido documentado que a condensação pode ocorrer por proteínas dirigentes (Davin and Lewis, 2000). Em quanto às peroxidases, se pode indicar que tem sido consideradas durante anos, as únicas classes de oxidases envolvidas na polimerização da lignina, novos resultados, a partir de sequenciamento de EST e estudos de expressão baseados em tecidos lignificados, tem amostrado que outras oxidases e particularmente as lacases, poderiam estar envolvidas na lignificação. Peroxidases e lacases fazem parte de uma grande família multigênica e tem sido dificultoso assinalar uma específica função para uma particular oxidase (Boudet, 2000). Além disso sinapil álcool é muito rápido oxidado em presença de *p*-coumarato, o qual é oxidado por peroxidases e transfere o radical para o álcool sinapil (Boudet, 2000). Dois peroxidases de milho (ZmPox2 e ZmPox3) tem sido localizadas em tecidos vasculares e em lignificação (de Obeso et al., 2003). Complementar a isto, ZmPox2, ZmPox3 e três novos genes de peroxidase do milho, ortólogos a genes de peroxidase de *Zinnia*, amostraram ser significativamente expressos em milho em internódios lignificando de plantas em etapa de espigamento (Guillaumie et al., 2007). No mesmo experimento, três genes de lacase, ortólogos aos genes de lacasse de álamo e a *zinnia*, foram simultaneamente expressos, um deles 2 vezes mais expresso do que o mais expresso para o caso do gene da peroxidase. Uma via menos direta para que os monolignóis entrem a ser parte do polímero de lignina, tem a ser possivelmente considerada, na ocorrência da coniferina (coniferil álcool 4-*O*- $\beta$ -D-glicosídeo), a qual se encontra como componente de reserva. Na primitiva árvore *Ginkgo biloba*, a coniferina foi mostrada a ser oxidada dentro de coniferaldeído glicosídeo por uma NADP-dependente de álcool desidrogenase, deglicosilada por uma  $\beta$ -glicosidase dentro de coniferaldeído e finalmente convertida dentro de coniferil álcool e incorporada dentro de lignina (Tsuji et al., 2005). Em magnólia (planta com reserva de coniferina), oleander ou eucalipto (plantas sem reserva de coniferina), a maioria da aglicona coniferina usada para a

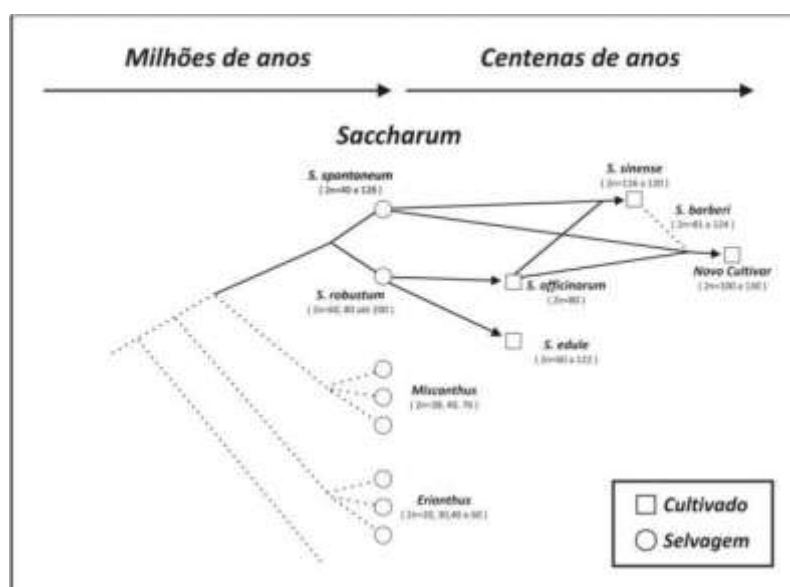
biossíntese da lignina via aldeído, foi incorporada maiormente em siringil, mas não como guaiacil lignina (Tsuji et al., 2004; Tsuji and Fukushima, 2004).

A recente identificação de outro gene não caracterizado da biossíntese da lignina, cafeoil shiquimato esterase (CSE), adiciona mais um passo, nesta rota metabólica que pode ser manipulado (Vanholme et al., 2013b). Estudos com mutantes T-DNA *cse* de *Arabidopsis* sugerem que a função da CSE na rota biosintética da lignina é a hidrólise do cafeoil shiquimato em cafeato. Vanholme et al. (2013b) trabalhando com mutantes *cse* de *Arabidopsis*, observou que estes, possuem níveis baixos de lignina e uma alta proporção de unidades H, associados com um aumento na eficiência de sacarificação de um 75% acima do tipo selvagem. Uns dos mutantes exibe uma baixa taxa de crescimento e uma reduzida produção de biomassa. Estes mesmos efeitos foram observados no trabalho de Ha et al. (2016), em mutantes de CSE para *Medicago truncatula*, indicando que CSE é crítica para uma normal lignificação nestas espécies. Segundo o trabalho de Ha et al. (2016), onde a partir de uma análise filogenética de 149 proteínas homologas de CSE correspondente a 61 espécies, usando a sequência CSE de *Arabidopsis* como isca, permitiu observar duas classes de CSE: I (*bona fide*, estreitamente relacionado à CSE de *Arabidopsis*) e II (com baixa identidade à CSE de *Arabidopsis*). *Panicum virgatum* possui dois homólogos agrupados na classe I. Sequências homologas de CSE de *Brachypodium distachyon* e *Zea mays* foram agrupadas na classe II. Em Ha et al. (2016) também se informa que a atividade CSE em extratos proteicos de caule de *P. virgatum* deu aproximadamente um 50 % de conversão de cafeoil shiquimato a cafeato como principal produto, e de forma oposta, iguais quantidades de extrato proteico de caule de *B. distachyon* e *Z. mays*, produziram muito pouca quantidade de cafeato. Não fica claro, se o baixo nível de formação de cafeato reflete a atividade de uma esterase não específica ou a atividade de um homólogo classe II, nessas espécies.

### 1.5. O gênero *Saccharum*

O gênero *Saccharum* engloba 6 espécies (Figura 2): *S. spontaneum*, *S. officinarum*, *S. robustum*, *S. edule*, *S. barberi* e *S. sinense* (D'Hont et al., 1998). Acredita-se que *S. officinarum* seja resultado de uma complexa introgressão entre *S. spontaneum*, *Erianthus arundinaceus* e *Miscanthus sinensis* (Daniels and Roach, 1987). *S. officinarum* é a única espécie dentro do gênero em que o número cromossômico não é variável entre os indivíduos (Price, 1963) e acredita-se sua origem foi a partir de *S. robustum*. *S. officinarum* tem um número de cromossomos de  $2n=80$  com um número básico de 10, fazendo isto uma

espécie poliplóide. Porém, *S. officinarum* não é um simples poliplóide, sendo um complexo híbrido de diferentes espécies e possui características de autopoliplóide (mais de dois conjuntos de cromossomos homólogos derivados somente de uma espécie) e também de alopólplóide (que possui dois ou mais cromossomos, mas diferentes) (Austrália, 2008). *S. spontaneum* é a espécie mais primitiva do gênero *Saccharum*, altamente polimórfica, resistente a doenças, estresses abióticos, vigorosa, com alto conteúdo de fibras, mas com muito pouco conteúdo de sacarose. Possui uma imensa variabilidade genética que pode contribuir com características de interesse agrônômico aos programas de melhoramento genético (Naidu and Sreenivasan, 1987), que visam o desenvolvimento de variedades comerciais com grande potencial para a produção de biomassa (Roach, 1978).



**Figura 2.** A provável história evolutiva das espécies de *Saccharum* (Grivet et al., 2006).

Diversas evidências moleculares indicam que *S. spontaneum* é muito diferente geneticamente em relação às outras espécies de *Saccharum* (Glaszmann et al., 1990; D’hont et al., 1993; Sobral et al., 1994; Burnquist et al., 1992; Vijayan Nair et al., 1999). *S. spontaneum* também é um complexo poliplóide com um provável número cromossômico básico de 8 ou 10 (D’Hont et al., 1998). Se acredita que *S. barberi* e *S. sinensis* são espécies selvagens, mas elas tem sido cultivadas desde tempos pré-históricos no norte da Índia e China, respectivamente. Esta situação envolveu um cruzamentos com outros gêneros e espécies e, conseqüentemente, *S. barberi* e *S. sinensis* são consideradas híbridos ancestrais intergenéricos (Daniels and Roach, 1987). Pensa-se que *S. barberi* é produto da introgressão entre *S.*



*officinarum* e *Erianthus*, quanto *S. sinensis* é da introgressão entre *S. officinarum* e *Miscanthus*. Cada uma das espécies *S. barberi* e *sinensis* contém os cromossomas homólogos de *S. officinarum* e *S. spontaneum*, bem como aos membros dos gêneros *Erianthus* e *Miscanthus*, indicando as origens complexas e inter-relações dentro do gênero *Saccharum* (D'Hont et al., 1996; Daniels and Roach, 1987). *S. barberi* possui um conteúdo moderado de sacarose, com resistência a estresses e alto conteúdo de fibras em relação a *S. officinarum*. Atualmente não existe grande interesse no seu melhoramento genético, principalmente devido à dificuldade de florescimento e à esterilidade. *S. robustum* é uma espécie selvagem, que possui colmos com alto conteúdo de fibras mas com níveis baixos de sacarose e se pensa que são os intermediários evolutivos entre *S. spontaneum* e *S. officinarum* (Austrália, 2008). Embora haja interesse potencial na utilização de *S. robustum* em programas de melhoramento genético por conta do seu grande vigor, há poucas citações da utilização da espécie, exceto nos híbridos produzidos pelo programa de melhoramento no Havaí (Naidu and Sreenivasan, 1987; Roach and Daniels, 1987). *S. edule* é morfologicamente similar a *S. robustum* exceto pela característica de que a ponta da flor ou inflorescência é compacta e é cultivada nas ilhas do Pacífico e Papua Nova Guiné, sendo também sugerido que *S. edule* é produto da introgressão de *S. officinarum* e *S. robustum* com outros gêneros (Daniels and Roach, 1987).

#### **1.6. Cana de açúcar e cana energia como fonte de biomassa lignocelulósica.**

Cana de açúcar é uma gramínea C4 perene, que cresce em climas tropicais e subtropicais (Moore, 1987; Somerville et al., 2010). A Cana de açúcar é uma matéria prima bem conhecida por acumular e armazenar altas concentrações de sacarose nos colmos, característica que é aproveitada para a alimentação é para a produção de E1G (Jung et al., 2012; de Souza et al., 2014). A grande maioria do açúcar no mundo é produzida a partir de cana de açúcar, o que a torna economicamente uma importante e valiosa cultura (Suprasanna et al., 2011).

O resíduo vegetal resultante do processo de extração de sacarose em cana é conhecido como bagaço, um material complexo, constituído principalmente por celulose (39%) hemicelulose (25%) e lignina (23 %) (Carroll and Somerville, 2009). O fato do bagaço ser um resíduo vegetal lignificado (Carroll and Somerville, 2009; Somerville et al., 2010) faz com que tenha potencial para ser usado como biocombustível, já que tem elevado poder calorífico (26MJ.kg-1) (Carvalho-Netto et al., 2014). Por esta razão é que atualmente o

bagaço vem sendo queimado para a geração de calor e eletricidade para uso na própria usina e/ou para a comercialização visando minimizar problemas com a crise no fornecimento de energia (Balbo and Padovani Neto, 1987). Atualmente, devido à demanda energética existe interesse em usar o bagaço para a produção de E2G, à partir do fracionamento da parede celular em açúcares fermentáveis, através de um processo chamado sacarificação. (Ruth and Thomas, 2003; Aden and Foust, 2009; Huang et al., 2009). O processo tem alto custo e é dificultado por uma série de razões, tais como: a celulose e a hemicelulose estão aglutinadas em um arranjo sistemático incrustado por lignina dificultando o acesso aos açúcares por enzimas hidrolíticas (recalcitrância), inibição do crescimento de microrganismos fermentadores pela liberação de compostos tóxicos provenientes da lignina. A maioria dos programas de melhoramento genético da cana de açúcar foram enfocados durante os últimos 100 anos no retrocruzamento dos híbridos de cana com as espécies ancestrais de *S. officinarum* (alto conteúdo de açúcar e baixo de fibras), para acrescentar a esta cultura a excepcional capacidade de desviar uma alta proporção de fotoassimilados em sacarose e armazená-la nos colmos (Carvalho-Netto et al., 2014), porém, isto reduz o vigor, e então as variedades híbridas são mais susceptíveis a estresses, doenças e consequentemente, a produtividade fica restrita (Waclawovsky et al., 2010). O potencial de produtividade de cana de açúcar é conhecido por ser quase 400 toneladas de biomassa fresca por ha/ano em condições ótimas, mas a real produtividade média é inferior a 100 toneladas por hectare (de Souza et al., 2014).

Mais recentemente tem sido dada atenção ao que se conhece por cana energia, uma forma distinta de cultura destinada para a produção de biomassa (fibra), ficando a produção de sacarose em segundo plano (Carvalho-Netto et al., 2014). A cana energia pode ser selecionada para muitos propósitos, desde a obtenção de calor, vapor e eletricidade, até várias formas de biocombustíveis (E1G e E2G). O melhoramento genético para o desenvolvimento da cana energia pode ser aproveitado a partir do potencial existente na ampla diversidade genética das espécies ancestrais do gênero *Saccharum* e outros gêneros relacionados, tais como *Erianthus*, *Miscanthus*, *Narenga* e *Sclerostachya*. As características essenciais para cana energia, tais como alta produtividade de biomassa, digestibilidade da parede celular, excepcional crescimento em áreas com pouca administração de nutrientes, perfilhamento e resistência a doenças, são encontrados em grande intensidade no germoplasma de *Saccharum* e espécies relacionadas (Bull and Glasziou, 1963). Recentes esforços do melhoramento estão tomando distintas vias, indicando que pode-se desenvolver

dois tipos de cana energia, com alta porcentagem de alelos da espécie *S. spontaneum* (alto porcentagem de fibra e baixo conteúdo de sacarose) (Ming et al., 2006). A cana energia de tipo I, que possui característica de cultura “dual”, com conteúdo de sacarose comparável às variedades híbridas produtores de sacarose, mas com muito mais fibra, e a cana de tipo II, desenvolvida apenas para a produção de biomassa, com alto teor de fibra, na qual a quantidade de sacarose é moderada (Tew and Cobill, 2008).

Entre as perspectiva futuras em relação à cana energia temos:

- Por causa da elevada produtividade de biomassa, a cana energia tem o potencial de produzir maiores quantidades de açúcares não celulósicos por hectare que as variedades híbridas de cana convencionais. Caso isto se cumpra, que a cana energia pode substituir, num futuro próximo, uma parte significativa das áreas destinadas para a cultura da cana de açúcar, mesmo aquelas destinadas para a indústria de E2G. Isso só será possível através da superação dos desafios no processamento da cana energia e o fornecimento de um conjunto de variedades especificamente desenvolvidos para este propósito, é dizer etanol de primeira ou segunda geração, vapor ou eletricidade e/ou açúcares celulósicos para a produção bioquímica.
- Por ser mais rústica, a cana energia pode ser plantada em terra com pouca necessidade de nutrientes, ou áreas de solo e climas piores do que aqueles reservados para a produção de alimentos ou cana convencional. A cultura de cana energia exige menor uso de fertilizantes, herbicidas e pesticidas.
- A cana energia pode produzir mais perfilhos e possui um coeficiente de multiplicação superior (1:30 ou mais) contra uma taxa menor (1:10) encontrada nas variedades híbridas atuais.
- A safra de cana energia será importante para a contenção da erosão do solo e na recuperação de solos degradados, pelo seu vigoroso sistema radicular na forma de rizoma, com raízes fortes e fasciculadas.

## 2. JUSTIFICATIVA

É conhecido que a cana de açúcar é uma das plantas mais eficientes em coletar e converter a luz solar em energia química, apresentando uma alta capacidade de acumular biomassa. Recentemente, com o advento do desenvolvimento da tecnologia do E2G, o bagaço (lignocelulose) tem recebido atenção como matéria prima para a produção de biocombustível, porém a biomassa de cana é lignificada (23%), característica que dificulta sua conversão em etanol, desde que a lignina limita a sacarificação por uma série de razões. Portanto, um estudo mais abrangente utilizando técnicas de biologia molecular e bioquímicas poderiam ampliar o conhecimento dos fatores que governam a recalcitrância conferida à parede celular pela lignina, a fim de aproveitar plenamente o potencial dos biocombustíveis celulósicos. Por outro lado, a base genética estreita encontrada nas variedades híbridas comerciais atuais limita os programas de melhoramento genético dirigidos a aumentar a qualidade da lignocelulose, situação que motiva o estudo dos genótipos ancestrais do gênero *Saccharum*: *S. spontaneum*, *S. barberi*, *S. robustum* e *S. officinarum*. Julga-se então, importante a caracterização do conteúdo, composição, estrutura e tipo de ligações das ligninas, bem como a identificação dos genes que participam no processo biossintético deste polímero, assim como a caracterização bioquímica da parede celular em espécies de *Saccharum*, para compreender o processo de lignificação.

### 3. OBJETIVOS

#### 3.1. Objetivo Geral

Estudar a biossíntese de lignina nas espécies *S. spontaneum*, *S. officinarum*, *S. robustum* e *S. barberi*, através da identificação de genes da via biossintética e da análise dos componentes da parede celular.

#### 3.2. Objetivos específicos

Utilizando colmos de diferentes idades (2<sup>o</sup>+3<sup>o</sup> colmos e 7<sup>o</sup>+8<sup>o</sup> colmos):

1. Estudar a composição da parede celular através de testes histoquímicos.
2. Determinar o conteúdo de lignina solúvel, insolúvel, de polissacarídeos estruturais (celulose, hemicelulose e pectina) e não estruturais (açúcares totais solúveis, sacarose, redutores e amido).
3. Determinar o porcentagem de monossacarídeos, ácidos hidroxicinâmicos e grupo O-acetil substituintes por espectroscopia 2D HSQC NMR.
4. Determinar o porcentagem de sacarificação
5. Determinar o conteúdo de fenóis totais.
6. Analisar o perfil de oligômeros da lignina e determinar a razão S/G.
7. Isolar e identificar genes envolvidos em cada passo enzimático da via biossintética da lignina

## 4. MATERIAIS E MÉTODOS

O presente trabalho de pesquisa se realizou durante os meses de fevereiro de 2014 a dezembro de 2015, no Departamento de Biologia Vegetal, do Instituto de Biologia da Unicamp.

### 4.1. Material Biológico e condições de crescimento

Plantas das espécies *S. spontaneum*, *S. officinarum*, *S. robustum* e *S. barberi*, provenientes de clones obtidos por cultura de tecidos fornecidos pela Dra Silvana Creste do Instituto Agronômico de Campinas, Ribeirão Preto, São Paulo, Brasil foram utilizados neste estudo. Os clones foram semeados em bandejas de plástico contendo vermiculita e mantidos em casa de vegetação, com irrigação diária e sem control de temperatura ou fotoperíodo. As plântulas crescidas foram transplantadas para vasos de 50 L (3 plantas por vaso) contendo substrato orgânico comercial (nitrogênio = 0,5%, umidade = 50%, carbono orgânico = 15%, pH = 6, C/N = 130 mmol/Kg, relação CIC/C = 8%, Genefertil-Genesol) e mantidos sob irrigação por aspersão (30 min, duas vezes por dia) em estufa por um período aproximadamente de um ano. Passado este período, os vasos foram transferidos para fora da estufa. O substrato dos vasos foi trocado e as plantas foram deixadas crescer por um período de mais 4 meses, sob irrigação por aspersão (30 min, três vezes por dia). A coleta do material foi feita em junho do 2014, sendo 5 repetições (5 vaso) por espécie. Colmos das repetições das diferentes espécies, aparentemente saudáveis, do mesmo estágio fisiológico e de tamanho uniforme foram coletados usando uma lamina de aço, separando os internódios 2°+3° (fase jovem) e 8° (fase madura), não incluindo o nó. Após serem lavados com água da torneira o material vegetal foi colocado em saco de polietileno e, em seguida, foram colocados em recipientes térmicos contendo nitrogênio líquido, para posterior liofilização. O material liofilizado foi macerado em mortar previamente esterilizados por autoclave e resfriado com nitrogênio líquido, e transferidos para tubos Falcon esterilizados, e armazenados em freezer a -80°C. Dessa forma, as amostras usadas nas análises eram provenientes de 4 espécies, dois estádios de desenvolvimento e em quintuplicatas biológicas, totalizando 40 amostras.

## 4.2. Análises histoquímicas

Internódios dos colmos 2+3 (fase jovem) e 8 (fase madura) dos caules das diferentes espécies foram utilizados nestas análises. Os testes histoquímicos foram realizados em secções feitas a mão livre, usando uma lamina de aço. Foram utilizados os seguintes reagentes para identificar os componentes da parede celular: a) Lignina: fluoroglucinol-HCL (Johansen, 1940), b) Monômeros siringil e guaiacil: reagente de Maüle (Mäule, 1901), c) Celulose: calcofluor white (Hughes and McCully, 1975), d) Amido: lugol (Berlyn and Miksche, 1976). Para a documentação dos resultados utilizou-se câmera de vídeo Olympus DP71 acoplada ao microscópio Olympus BX 51.

## 4.3. Análises bioquímicas

**Determinação dos polissacarídeos da parede celular:** As determinações de celulose, hemicelulose e pectina foram realizadas de acordo com a metodologia descrita por (Chen et al., 2002) com algumas modificações: 150 mg de amostra foi homogeneizada com 5 ml H<sub>2</sub>O destilada fria, usando um vortex e, em seguida, foi centrifugada durante 10 min, removendo-se o sobrenadante e conservando-se o pellet. O pellet foi lavado sequencialmente com 5 ml de H<sub>2</sub>O destilada fria, 5 ml de acetona e 5 ml de metanol/clorofórmio (1:1, v/v), sendo que em cada passo foi feita a centrifugação para recuperação do material sólido, que foi então seco em bancada por 24 h antes de ser extraído com o solvente seguinte. Para remover o amido, o resíduo obtido foi tratado a 37°C durante 3 h com 5 ml de 2 unidades.ml<sup>-1</sup> de amilase pancreática (tipo I-A, Sigma Chemical Co, St Louis, MO) em tampão 0,1 M acetato de sódio (pH: 6,5), e em seguida centrifugado para recuperar o material sólido. Este foi então extraído três vezes com 3 ml de 20 mM de oxalato de amônio (pH 4,0) a 70°C durante 1 h para se obter a fração correspondente às pectinas. Após centrifugação, o pellet foi então extraído com 3 ml de NaOH 0,1 M à temperatura ambiente sob vácuo durante 24 h no escuro, e o material foi recuperado por centrifugação. O sobrenadante foi denominado fração 1 de hemicelulose (HEM1). O pellet foi extraído três vezes com 2 ml de NaOH a 17,5% por 8 h, sob a vácuo e os sobrenadantes obtidos de cada extração foram reunidos e chamados fração 2 de hemicelulose (HEM2). A fração total de hemicelulose foi obtida pela soma das extrações (HEM1 + HEM 2). O resíduo alcalino insolúvel foi lavado sequencialmente com 5 ml de água destilada, 5 ml de 1 mM de ácido acético e 5 ml de etanol e em seguida seco a 37°C. Em

seguida, esta fração foi dissolvida em 5 ml de uma solução 72% de H<sub>2</sub>SO<sub>4</sub> durante 1 h à temperatura ambiente e diluída 30 vezes com água destilada para obter a fração de celulose correspondente. O teor total de açúcares totais em cada fração foi obtido de acordo com o método do fenol sulfúrico, utilizando glucose como padrão (DuBois et al., 1956).

**Determinação de açúcares solúveis totais, sacarose, açúcares redutores:** Para a determinação dos carboidratos não estruturais, se pesou 20 mg de tecido liofilizado em tubos eppendorf, colocando logo depois 1,5 ml de etanol 70%. Após a incubação em banho-maria a 60°C durante 30 min, as amostras foram centrifugadas a 6.000 xg por 10 min. O sobrenadante foi coletado e o precipitado solubilizado novamente em 1,5 ml de etanol 70%, sendo re-extraído como mencionado anteriormente, por mais duas vezes. As três alíquotas coletadas de cada amostra foram reunidas, perfazendo um total de 4,5 ml de extrato bruto, material que serviu para as análises de açúcares solúveis, sacarose e açúcares redutores. Para a determinação de Açúcares solúveis totais y sacarose foi usado o método do fenol sulfúrico (DuBois et al., 1956) mediante a leitura no espectrofotômetro a 495 nm, usando uma curva padrão de glicose no intervalo 20-100 ug/ml e sacarose, respectivamente, indicando que para a análise quantitativa de sacarose a amostra foi fervida por 10 min com KOH 30% (van Handel, 1968). As análises para determinar o conteúdo de açúcares redutores foram feitas de acordo a metodologia descrita por (Somogyi, 1937; Nelson, 1944), usando uma curva padrão de glicose no intervalo 10-50 ug/ml.

**Análise de constituintes da parede celular por espectroscopia 2D HSQC NMR:** Estas análises foram feitas no laboratório do Dr. Markus Pauly, na California University em Berkeley. 25 mg de amostra vegetais de cada espécie de *Saccharum* foram submetidas a lavagens sucessivas com água destilada quente usando funil de Buchner e depois em álcool. Em seguida foram extraídas com 0, 75 ml de DMSO-d<sub>6</sub> e 10 µl de [Emim] oAc-d<sub>14</sub>, de acordo, com a metodologia descrita por (Cheng et al., 2013; Chong et al., 2014). Posteriormente, as amostras foram submetidas a uma avaliação por 2D HSQC NMR em um espectrômetro Bruker AVANCE 600 Mhz NMR, equipado com um crio-sonda de 5 mm TXI <sup>1</sup>H/<sup>13</sup>C/<sup>15</sup>N, usando a sequência pulso ‘hsqcetgpsisp.2’. Os experimentos foram realizados a 25°C, usando os seguintes parâmetros: largura espectral de 12 ppm na dimensão F2(<sup>1</sup>H) com 4096 pontos (TD1) e 160 ppm na dimensão F1(<sup>13</sup>C) com 256 pontos (TD2); número de avaliações (SN) de 200 e atraso entre as avaliações (D1) de 1 segundo. Os desvios químicos foram referenciados aos picos do solvente DMSO (δ<sub>C</sub> 39.5 ppm, δ<sub>H</sub> 2.5 ppm). Os dados de NMR foram quantificados usando o software Bruker’s Topspin 3.1, de acordo com a



metodologia de (Cheng et al. 2013). A acetilação dos xilanos foi quantificada como brevemente se segue: os sinais na região aromática (sinais H1-C1 de 2-*O*-Ac-Xyl, 3- *O*-Ac-Xyl, 2,3-*O*-Ac-Xyl, Xyl (xilano) e os extremos redutores de xilanos ( $\alpha/\beta$ -Xyl-R)), foram somados até o 100%, e os sinais na região alifática foram integrados separadamente para calcular o conteúdo relativo de cada forma das unidades *O*-acetil-xilano. O conteúdo relativo das unidades 2-*O*-acetil e 2,3-*O*-Acetil-xilano foram calculadas da sinal H2-C2 e as unidades 3-*O*-Ac-xilano foram calculadas das sinal H3-C3. Os monossacarídeos (glicose (Glu), xilose (Xyl), arabinose (Ara) e manose (Man)) da hemicelulose foram quantificados das integrais anoméricas como uma fração do 100%. FA (ferulato) e *p*CA (*p*-coumarato) foram quantificados das suas integrais aromáticas de lignina como uma fração do 100%.

**Determinação de amido:** A determinação de amido foi realizada de acordo com a metodologia descrita por (Amaral et al., 2007) com pequenas modificações: Em eppendorf contendo 10 mg de amostra foram realizadas quatro extrações com 500  $\mu$ L de etanol 80 % a 80°C (em banho-maria) por 20 min, para retirada de açúcares, pigmentos, fenóis e outras substâncias. O precipitado foi então seco a temperatura ambiente até completa evaporação do resíduo de etanol. Em seguida foram adicionados 500  $\mu$ L (120U  $\text{ml}^{-1}$ ) de  $\alpha$ -amilase (EC 3.2.1.1) termoestável de *Bacillus licheniformis* (cód. E-ANAAM, MEGAZYME, Irlanda), diluída em tampão MOPS 10 mM e pH 6,5. As amostras foram incubadas a 75°C (em banho-maria) por 30 min. Este passo foi repetido por mais uma vez, totalizando 120 unidades de enzima. Em seguida, adicionou-se 500  $\mu$ L (30U  $\text{ml}^{-1}$ ) de amiloglucosidase (EC 3.2.1.3) de *Aspergillus Níger* (cód. E-AMGPU, MEGAZYME, Irlanda) em tampão acetato de sódio 100 mM e pH 4,5. As amostras foram incubadas a 50°C (em banho-maria) por 30 min. Este passo foi repetido por mais uma vez, totalizando 30 unidades de enzima. Foram acrescentados então, 100  $\mu$ L de  $\text{HClO}_4$  0,8 M para parar a reação e precipitar proteínas. A dosagem de amido procedeu-se da seguinte forma: em eppendorf contendo 10  $\mu$ L de extrato foi adicionado 40  $\mu$ L de água destilada e 750  $\mu$ L do reagente glicose PAP Liquiform (Labtest Diagnóstica S.A.). A mistura foi levada ao banho-maria a 37°C por 15 min. O teor de glicose foi determinado em leitor de microplacas de ELISA (modelo EL307C, Bio-Tek Instruments, Winooski, Vermont) em comprimento de onda de 490 nm. As concentrações de glicose foram calculadas a partir de uma curva-padrão obtida com leituras de soluções contendo 5, 10, 15, 25 e 30  $\mu$ g de glicose. Como branco, colocou-se água no lugar da amostra.

**Determinação de compostos fenólicos solúveis totais:** O conteúdo total de compostos fenólicos solúveis foi determinado com o reagente de Folin-Ciocalteu (Swain and Hillis, 1959), utilizando-se ácido clorogênico como padrão. Amostras de 150 mg foram colocadas em eppendorf de 1,5 mL e extraídas com 700  $\mu$ L de etanol 80% sob sonicação por 30 min, sendo depois centrifugadas durante 10 min a 12000 rpm. Coletou-se o sobrenadante (por volta de 600  $\mu$ L) e se repetiu a extração anterior. Os extratos de cada amostra foram juntados e 50  $\mu$ L foram adequadamente diluídos com 3,450 ml de água destilada. Em seguida, foram adicionados 0.25 ml do reagente de Folin-Ciocalteu, misturando a solução e incubando-a por 3 min em temperatura ambiente. Logo depois, 0.5 ml de  $\text{Na}_2\text{CO}_3$  saturado (17,5g/50 ml) foram misturados à solução, agitados e mantidos em repouso por 1 h em temperatura ambiente, quando então se adicionou 0.75 ml de água destilada. A determinação do conteúdo de fenóis totais foi feita mediante a leitura da absorbância em 725 nm num espectrofotômetro UV-VIS. Se utilizou uma curva padrão de ácido clorogênico calculada para o intervalo de concentração de 100-1000  $\mu$ g/ml.

**Determinação de lignina solúvel e insolúvel-Klason:** A determinação de lignina solúvel e insolúvel foi conduzida de acordo com o protocolo TAPPI UM-250 (TAPPI, 1985) com pequenas modificações. Amostras de 200 mg foram pesadas e colocados em saquinhos de TNT (tecido-não-tecido) para serem posteriormente transferidas para extrator Soxhlets, sendo lavados com 100 ml de acetona, mínimo para 8 horas para remover os componentes que podem ser "pseudolignina" durante a análise de Klason. O material lavado ficou secando durante pelo menos 24 h à temperatura ambiente em capela, sendo então pesado e transferido para tubos de ensaio (com tampa rosqueada e Teflon) previamente resfriados em gelo, contendo 3 ml de  $\text{H}_2\text{SO}_4$  72%, misturando o conteúdo em vortex durante 1 min. Subsequentemente os tubos foram colocados em banho-maria a 20°C, agitando cada 10 min em vortex durante 2 h. Após o período de hidrólise, o conteúdo de cada tubo foi transferido para frascos de vidro, utilizando 112 ml de água Milli-Q para lavar resíduos nos tubos de hidrólise. Os frascos contendo 115 ml de 4% de  $\text{H}_2\text{SO}_4$ , mais os materiais hidrolisados foram selados e autoclavados a 121 ° C durante 60 min, após este período, os frascos que contêm as amostras foram resfriados até temperatura ambiente e os hidrolisados foram filtrados em vácuo em filtros de microfibras de vidro Whatman GF / C 47 mm (estes filtros foram colocados durante 24 horas numa estufa a 105 ° C), lavando-se estes com 200 ml de água

Milli-Q pré-aquecida a 50 ° C para remover os resíduos de ácidos e açúcares. O filtrado obtido foi usado na determinação de lignina solúvel, lendo em UV-VIS Ultrospec Pharmacia Biotech 1000 a 205 nm, utilizando o coeficiente de extinção de 110 l. g<sup>-1</sup>. cm<sup>-1</sup>. O teor de lignina solúveis e insolúveis se expressara em mg/g massa seca. O conteúdo de lignina insolúvel é determinado gravimetricamente pela diferença de peso do filtro de microfibras de vidro de filtro antes e após este processo.

**Determinação da razão S/G:** A determinação da razão S/G foi feita de acordo com (Mokochinski et al., 2015). Amostras de 100 mg foram hidrolisadas com uma solução de 2,0 ml de NaOH 4 M em tubos de ensaio com tampa de rosca e vedados com fita teflon. Os tubos foram aquecidos durante 24 h a 95°C num sistema de aquecimento do tipo *dry block*. Os tubos foram resfriados em temperatura ambiente e as amostras foram acidificadas com aproximadamente 1,6 ml de HCl 6 M. O tubo foi agitado para garantir a neutralização da base. Após a acidificação, o pH foi conferido com fita de papel indicador. Em seguida a amostra foi centrifugada a 13000 rpm por 5 min e uma alíquota de 500 µL do sobrenadante foi transferida para outro recipiente (eppendorf de 2,0 mL) e extraída duas vezes com 1,0 mL de acetato de etila, juntando as duas frações da fase orgânica. O solvente foi evaporado sob fluxo de N<sub>2</sub> e o conteúdo foi dissolvido em 1,0 mL de água Milli-Q. Em seguida as amostras foram analisadas por UHPLC-MS. Se utilizou um sistema de UHPLC acoplado com um espectrômetro de massas do tipo triplo quadruplo equipado com fonte de ionização ESI (modelo ACQUITY, Waters Corp, Manchester, UK). O software MassLynx V.4.1 (Waters Corp. Manchester, UK) foi utilizado para aquisição e tratamento dos dados. Os compostos provenientes da hidrólise alcalina foram separados usando uma coluna C<sub>8</sub> (50 mm × 2,1 mm; 1,7 µm). Como fase móvel se usará água (A) e acetonitrila (B), seguindo o gradiente: 90% A até 4,5 minutos, mudando para 100% B em 8,0 minutos. Essa condição foi mantida constante até 9,0 minutos, e em seguida se retornou para condição inicial estabilizando em 10,0 minutos. O fluxo se manteve constante durante a corrida em 0,300 mL/min. O volume de injeção foi de 5,0 µL. Os espectros de massas foram adquiridos em modo SIM para os íons *m/z* 121, 151 e 181, em modo negativo sendo a voltagem do capilar -3,0 KV e do cone -52, -26 e -30 V para H, G e S, respectivamente. A temperatura da fonte e de dessolvatação foi de 150 e 350 °C, respectivamente.

**Determinação do porcentagem de sacarificação:** A determinação do porcentagem de sacarificação foi conduzida de acordo com a metodologia descrita por (Brown and Torget, 1996), com pequenas modificações. Em tubos eppendorfs, foi pesada uma quantidade de biomassa liofilizada equivalente a 10 mg de celulose (previamente determinada, conforme protocolo acima e livre de açúcares solúveis e amido), e a esta foi adicionado 500 µl de tampão citrato de sódio (0,1 M, pH 4,8), 10 µl de azida de sódio e uma quantidade necessária de água para completar 1 ml. Posteriormente, a mistura foi aquecida a 50°C e em seguida acrescentou-se 6,08 µl de uma mistura (1:4, v/v) de enzimas celulase (1,2 FPU/10 mg de celulose) e celobiose (1,26 U *p*NPGU/10mg celulose). Este coquetel foi preparado a partir de celobiose de *Aspergillus niger* e celulase de *Trichoderma reesei* (Sigma-Aldrich). Os eppendorfs contendo as amostras e a solução de reação, foram fechados hermeticamente e colocados inclinados (45°) num shaker, a 50°C, sendo mantidos a 160 rpm durante 5 dias. Passado esse período as amostras foram centrifugadas a 12000 rpm por 15 min. A determinação de glicose foi feita no sobrenadante pelo método do fenol sulfúrico (DuBois et al., 1956) mediante a leitura no espectrofotômetro a 495 nm, usando uma curva padrão de glicose no intervalo 20-100 µg/ml.

**Análise de oligômeros de lignina por LC-MS:** A Análise de oligômeros de lignina foi realizada de acordo com a metodologia descrita por (Kiyota et al., 2012). 100 mg de material previamente liofilizado e macerado foram transferidos para eppendorfs de 1,5 mL, e foram extraídos com 700 µL de etanol 80% sob sonicação por 30 min. Em seguida os tubos foram centrifugados durante 10 min a 12000 rpm e coletou-se o sobrenadante (por volta de 600 µL), que foi transferido para um tubo eppendorf de 2 mL. Se repetiu a extração anterior e os extratos foram juntados para em seguida serem secos Speed-Vac. A partir deste ponto pode-se guardar o material em freezer -20°C. O material seco foi solubilizado em 300 µL de água Milli-Q e 300 µL de acetato de etila. Agitou-se em vortex por alguns segundos e os tubos permaneceram em descanso para que a fase orgânica (superior) se separasse da fase aquosa (inferior). A fase orgânica (com os oligômeros das ligninas) foi coletada em um novo tubo e deixada secar sob o fluxo da capela. O material foi solubilizado em 300 µL de acetonitrila/água (1:2, V/V) logo antes das análises. 10 µL de amostra foram injetados num Acquity UPLC acoplado a um espectrômetro de massas TQD triplo quadrupolo (Micromass-Waters, Manchester, England). A separação cromatográfica foi feita usando uma coluna

Waters Acquity C18-BEH (2.1 x 50 mm, 1.7  $\mu$ m) com eluição num gradiente de 5 a 100% de acetonitrila (Solvente B) em 8 min, e usando água Milli-Q com 0.1% de ácido fórmico como solvente A. Um fluxo de 0.200 ml.min<sup>-1</sup> e temperatura de coluna de 30°C foram usados. A ionização em eletrospray se deu em modo negativo: capilar 3.0 kV, cone 50 V temperatura na fonte de ionização 150°C e temperatura de dessolvatação 300°C. Os espectros MS/MS foram obtidos por dissociação por colisão induzida (CID) com energia de colisão de 20V.

#### 4.4. Análises moleculares

**Análise *in silico* de bancos de dados e desenho de primers:** Foram buscadas sequências para os genes das seguintes enzimas da biossíntese da lignina: 4-hidroxicinamoil CoA ligase (4CL; EC 6.2.1.12), cinamoil-CoA redutase (CCR; EC 1.2.1.44), ferulato 5-hidroxilase (F5H; EC 1.14.13.-), cafeato *O*-metiltransferase (COMT; EC 2.1.1.68), cinamil álcool desidrogenase (CAD; EC 1.1.1.195), fenilalanina amônia-liase (PAL; EC 4.3.1.5), cafeoil CoA 3-*O*-metiltransferase (CCoAOMT, EC 2.1.1.104), *p*-coumaroil shiquimate 3'-hidroxilase (C3'H, EC 1.14.13.36), cinamato 4-hidroxilase (C4H, EC 1.14.13.11), hidroxicinamoil-coenzima A chiquimate:quinato hidroxicinamoiltransferase (HCT; EC 2.3.1.133). Usando como iscas as sequências dos genes mais expressos nos tecidos de variedades híbridas de cana identificadas por (Bottcher et al., 2013), foram feitas buscas nos bancos de dados NCBI, Gene Index e Phytozome, sendo que entre as sequências que retornarem foram selecionadas as sequências de sorgo (*Sorghum bicolor*), arroz (*Oryza sativa*), milho (*Zea mays*), trigo (*Triticum aestivum*), azevém (*Lolium perenne*) e *Arabidopsis thaliana*. Somente foram utilizadas sequências com CDS completos e de menor “e-value”. Estas sequências foram alinhadas com o uso do programa BioEDIT (Hall, 1999) e foram desenhados primers em regiões conservadas. Quando isto não foi possível, foram desenhados primers degenerados. Os primers utilizados são mostrados na Tabela 1 e suas sequências foram obtidas usando programa Primer 3. Procurou-se posicionar os primers de tal forma que cobrissem praticamente toda a ORF. Para 4CL, C3H, C4H e HCT, com um “open reading frame - ORF” acima de mil bases o sequenciamento foi feito em duas etapas, sendo feitos pares de primers internos, para poder dar total cobertura da sequência do gene.

**Tabela 1.** Sequências dos pares de primers utilizados na amplificação dos genes da rota biossintética da lignina em espécies de *Saccharum*.

Enzima	Forward 5'-->3'	Reverse 5'-->3'	Amplicon (pb)
<b>4CL</b>	AGCTKCCGGACATCGASATC	CTTGTAGAAMACCACCTCCTTKG	1340
	GCCGCTGTTCCACATCTACT	-----	550
	CCCTCTCGCATGTACAGGTT	-----	550
<b>CAD</b>	TACTGYGGSATATGCCAC	CGAYGTCGATGACGAAGC	973
<b>C3H</b>	AACCTGCGCCAGATCAAG	GTGCCCATGAASGTGACG	1343
	ACCTTTCTATGGTGGCCTTC	-----	550
<b>C4H</b>	GTTCGGCGACATCTTCCTC	CTTCTCCGTGGTGTCGATCT	1281
<b>CCoAOMT</b>	CTCAAGAGCGAAGMCCTSTAC	GAGCTGGCAGAYCTCGAC	649
<b>CCR</b>	GSCTCGTGGSTCGTCAAG	GGTTSGAGAACTTGTACGGCT	823
<b>COMT</b>	GTCCATCCTGCCCATGAC	CAGTCGTGGAGGATCCACTT	756
<b>HCT</b>	GGGGTCGGAGATGGTGTA	CGGAACTTCTCCATGTGCTC	1317
	ATCGACTACTTCGGCGACTTC	-----	550
<b>F5H</b>	CGWTGATGGACCAGCTGAC	TGCTCGTCGATGATCTTGTC	625

**Extração de RNA total e a síntese da primeira fita de cDNA:** Num eppendorf de 2 mL foi adicionado aproximadamente 1/3 do volume com o pó de uma mistura 1/1 de tecidos de internódios jovens e maduros ((2+3°) e 8° internódios), previamente macerados em nitrogênio líquido. A extração de RNA total foi feita pelo método de Trizol (Reagente Tri-Phasis - BioAgency) descrito por (Porto et al., 2011). A quantificação do RNA total foi feita em espectrofotômetro a 260 nm e pureza pela razão 260 nm/280 nm. A qualidade do RNA foi verificada por eletroforese em gel agarose 1% com brometo de etídio e visualizada sob luz UV em foto-documentador (Gel Doc 2000, BIO RAD). As amostras de RNA total foram tratadas usando o kit Turbo DNase-free (Ambion) e submetidas à transcrição reversa usando o kit SuperScript III de síntese de cDNA primeira fita, conforme instruções do fabricante (Invitrogen).

**Amplificação, clonagem e sequenciamento:** O cDNA da amostra combinada de colmos jovens e maduros foi utilizado para amplificação de fragmentos dos genes da biossíntese da lignina usando os primers desenhados para todos os genes de cada enzima da rota biossintética da lignina. Foram feitas reações de RT-PCR usando um Termociclador Veriti®

96-Well Thermal Cycler-AB Applied Biosystems. As condições para a amplificação dos genes foram 40 ciclos, consistindo cada ciclo de uma desnaturação a 94°C por 30 s, anelamento a 57°C por 1 min, extensão a 72°C por 2 min, com um passo inicial de desnaturação de 94°C por 5 min e uma extensão final de 72°C por 5 min. As concentrações finais para 25 µL de reação de PCR foram: 200 µM de dNTP, 1,5 mM de MgCl<sub>2</sub>, 1 U de Taq-Polimerase em Buffer 1X Quiagen (Tris-Cl, KCl, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, pH 8.7) 0,4 µM de cada primer, concentração variável de cDNA, DMSO, Trehalose, BSA e Formamida. As bandas amplificadas foram separadas em gel de agarose 1% contendo brometo de etídeo e visualizadas em um foto-documentador Gel-Doc 2000 Biorad. Elas foram recuperadas do gel usando kit GeneJET Extraction- Thermo Scientific® e inseridas no vetor de clonagem “pGEM®-T easy Vector System” (Promega), conforme as instruções do fabricante. Os vetores foram clonados em *Escherichia coli* DH10β termo-competentes (Novagen), seguido de plaqueamento em meio LB semi-sólido contendo ampicilina, IPTG e XGal, e incubadas a 37°C *overnight*. As colônias transformadas de coloração branca foram selecionadas para em seguida serem analisadas por “miniprep” pelo método de lise alcalina (*PureLink Quick Plasmid Miniprep Kit*, Invitrogen) para verificar a presença do inserto. A presença do inserto foi confirmada por uma reação de restrição com EcoRI (Promega), que corta o vetor pGEM-T Easy em ambas as extremidades, liberando o fragmento inserido, que foi visualizado em gel de agarose. Para cada enzima foram verificadas tantas colônias sejam necessárias até que os “minipreps” identifiquem 25 colônias com insertos. Todas as 25 colônias de cada enzima foram enviadas para sequenciamento usando os *primers* flanqueadores do vetor (Forward e Reverse M13). O sequenciamento foi feito pelo Laboratório Central de Tecnologias de Alto Desempenho em Ciências da Vida (LaCTAD)

#### 4.5. Análises filogenéticas

A sequência dos transcritos dos genes da rota biossintética da lignina nas espécies de *Saccharum*, obtidos a partir do sequenciamento, foram traduzidas para aminoácidos e foram feitas buscas de proteínas homologas nos bancos de dados NCBI (<http://www.ncbi.nlm.nih.gov/>), Gene Index (<http://compbio.dfci.harvard.edu/tgi/plant.html>), SUCEST (<http://sucest-fun.org>) e Phytozome (<http://www.phytozome.net/>) para construir árvores filogenéticas. O alinhamento múltiplo de sequências de aminoácidos foi realizado com o programa ClustalW. A árvore filogenética foi construída usando o programa MEGA

versão 4.02 e as relações evolutivas foram inferidas usando o algoritmo Neighbor-joining com Bootstrap para 1000 repetições. “Gap regions” foram excluídas por ajuste manual.

#### **4.6. Análises estatísticas**

Se aplicou um desenho completamente randomizado 4x2, onde o primeiro nível é as espécies e no segundo nível foram os estádios de maturação da cana (internódios maduro e jovem). Foi realizada uma análise de variância (ANOVA) e quando uma variação significativa foi encontrada a comparação de médias foi feita pelo teste de Tuckey ( $p < 0,05$ ), usando o pacote estatístico Statistica 8.0. Para a maioria das análises bioquímicas foram feitas 5 quintuplicatas biológicas com três replicatas técnicas. Para as avaliações da relação S/G, oligômeros de lignina, ácidos hidroxicinâmicos, monossacarídeos e xilanos acetilados foram feitas uma replicata técnica. Os resultados das análises foram expressados pela média e erro padrão ( $n=5$ ).



## 5. RESULTADOS

### 5.1. Carboidratos estruturais da parede celular

Independentemente do estágio de desenvolvimento, o conteúdo de celulose foi maior em *S. spontaneum* e *S. robustum* e menor nas espécies *S. officinarum* e *S. barberi* (Figura 3A). Nas duas primeiras espécies, o maior conteúdo foi em colmos do internódio 8. Quanto à hemicelulose, o teor sempre foi maior no internódios 2/3 (Figura 3B) sendo que a espécie com menor teor foi *S. Officinarum* para os internódio jovens e maduros. As outras espécie apresentaram valores semelhantes entre elas para os dois estágios de desenvolvimento. Os teores de pectina (Figura 3C) foi maior nos internódios novos de *S. barberi* e *S. officinarum* comparados ao maduro e semelhantes entre els nas outras duas espécies. O maior teor de pectina foi encontrado no internódios 2/3 de *S. officinarum* e o maior teor no internódio 8 foi encontrado em *S. spontaneum*.

### 5.2. Carboidratos não estruturais

Os maiores teores de sacarose (Figura 4B), o açúcar que mais se acumula nos colmos de cana de açúcar, e por consequência de açúcares solúveis totais (Figura 4A), foram encontrados nos colmos maduros de *S. officinarum* e *S. barberi*. Estas duas espécies foram as que acumularam mais açúcares redutores nos colmos novos (Figura 4C) e por consequência, também foram as que apresentaram maior teor de açúcares solúveis nos colmos. Enquanto os teores de sacarose foram semelhantes nos colmos novos e maduros de *S. robustum* e *S. spontaneum*, os teores de redutores nestas espécies foram maiores nos colmos novos. Inesperadamente, o teor de amido em *S. spontaneum* foi mais que oito vezes maior que nas outras espécies para os colmos maduros (Figura 4D). Comparativamente, colmos novos de *S. officinarum* acumularam mais amido que colmos maduros.

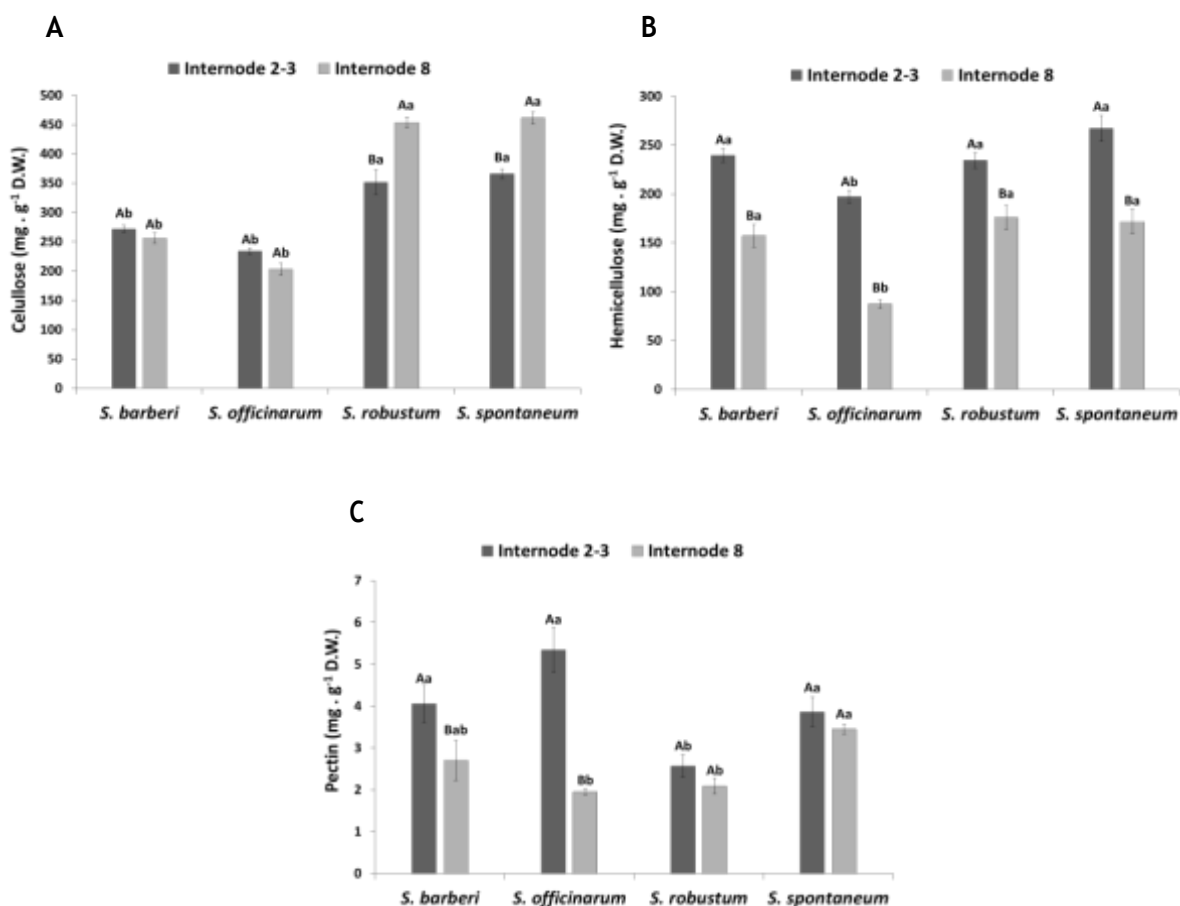


Figura 3. Conteúdo de (A) celulose (B) hemicelulose e (C) pectina nos internódios de espécies de *Saccharum*. Letras maiúsculas diferentes denotam diferenças significativas entre internódios de diferente estágio de desenvolvimento dentro de uma mesma espécie. Letras minúsculas diferentes indicam diferenças entre internódios de um mesmo estágio de desenvolvimento das diferentes espécies. As médias foram comparadas pelo teste de Tukey,  $P \leq 0,05$ . As barras verticais indicam o erro padrão das médias de quintuplicatas independentes.

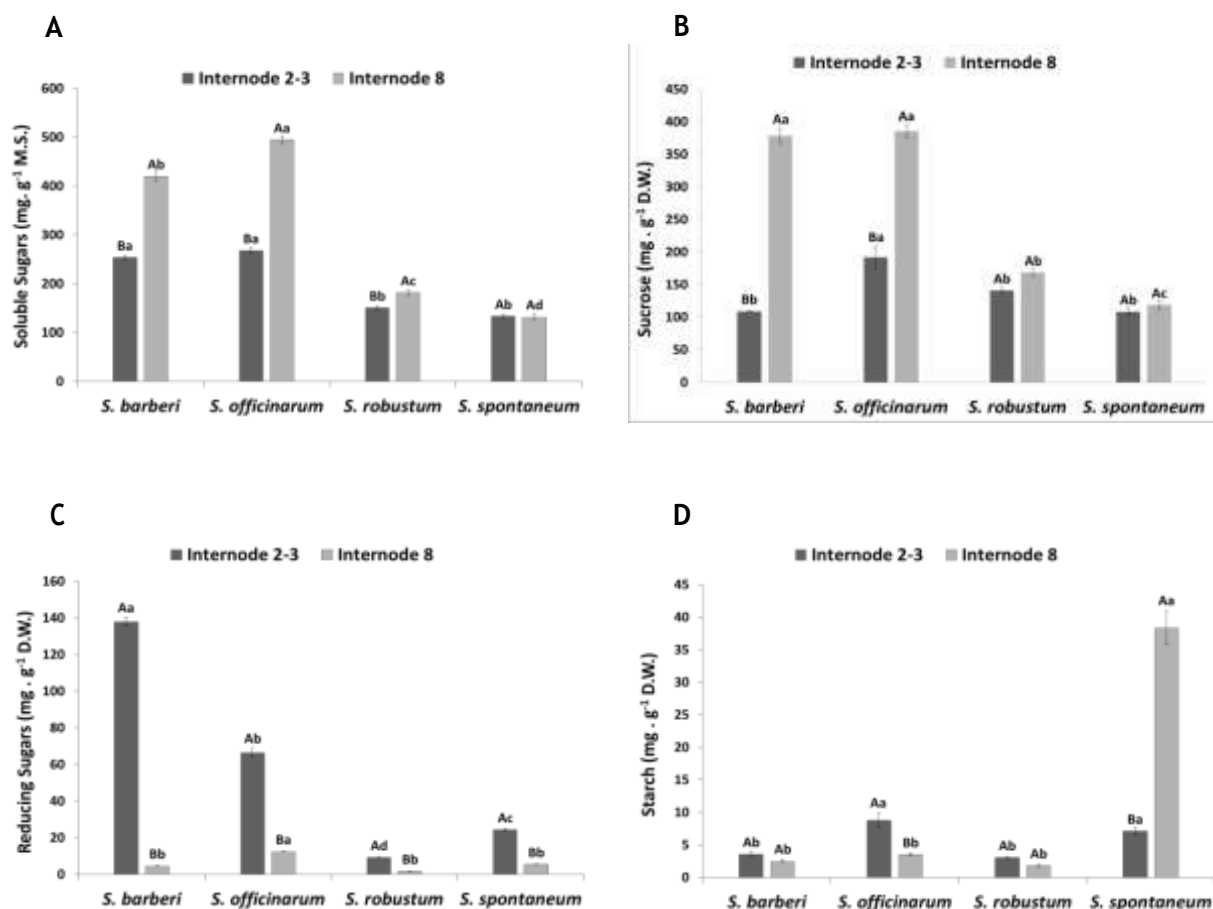


Figura 4. Conteúdo de (A) açúcares solúveis totais, (B) sacarose, (C) açúcares redutores e (D) amido nos internódios de espécies de *Saccharum*. Letras maiúsculas diferentes denotam diferenças significantes entre internódios de diferente estágio de desenvolvimento dentro de uma mesma espécie. Letras minúsculas diferentes indicam diferenças significantes entre internódios de um mesmo estágio de desenvolvimento das diferentes espécies. As médias foram comparadas pelo teste de Tukey,  $P \leq 0,05$ . As barras verticais indicam o erro padrão das médias de quintuplicatas independentes.

### 5.3. Lignina solúvel e insolúvel

Nas quatro espécies os maiores teores de lignina solúvel foram encontrados nos internódios novos (Figura 5A), sendo que *S. officinarum* teve o maior teor, que foi igual nas outras três espécies. *S. officinarum* também mostrou o maior teor nos colmos maduros. Por outro lado, o teor de lignina insolúvel foi maior no internódio maduro nas quatro espécies (Figura 5B), sendo que nos dois estádios os maiores teores foram encontrados em *S. spontaneum* e *S. robustum*.

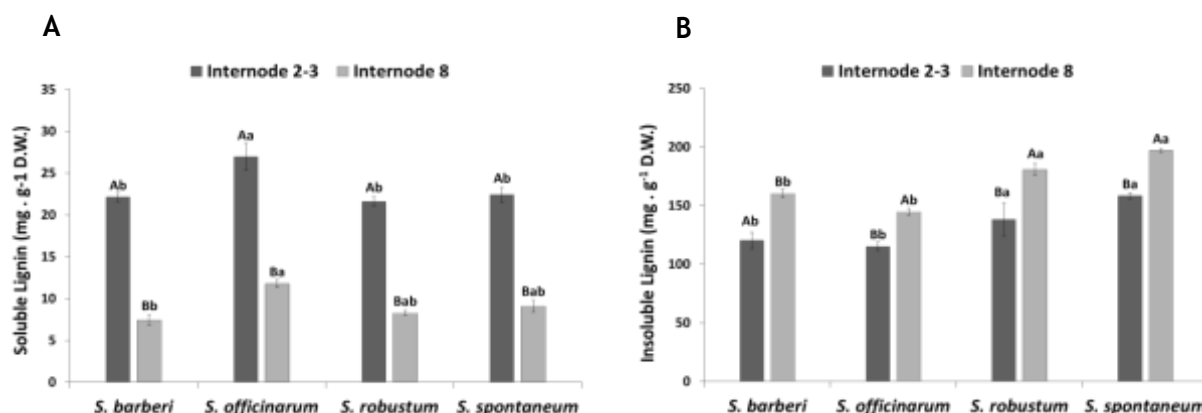


Figura 5. Conteúdo de (A) lignina solúvel e (B) insolúvel nos internódios de espécies de *Saccharum*. Letras maiúsculas diferentes denotam diferenças significantes entre internódios de diferente estágio de desenvolvimento dentro de uma mesma espécie. Letras minúsculas diferentes indicam diferenças significantes entre internódios de um mesmo estágio de desenvolvimento das diferentes espécies. As médias foram comparadas pelo teste de Tukey,  $P \leq 0,05$ . As barras verticais indicam o erro padrão das médias de quintuplicatas independentes.

#### 5.4. Porcentagem de sacarificação

Houve um padrão semelhante na porcentagem de sacarificação entre *S. barberi* e *S. officinarum*, e entre *S. robustum* e *S. spontaneum*, formando dois grupos distintos. Enquanto que nas duas primeiras espécies a sacarificações entre os dois estágios de internódios foram próximas, elas foram bastante diferentes no segundo grupo. A sacarificação nos internódios maduros de *S. spontaneum* e *S. robustum* foi praticamente metade dos valores encontrados com os internódios jovens. De modo geral, o percentual de sacarificação nos colmos jovens foram próximos nas quatro espécies, ficando ao redor de 65%.

#### 5.5. Razão S/G da lignina

A razão S/G foi mai alta nos internódios novos do que nos maduros de *S. robustum* e *S. barberi*, mas não diferiram nas outras duas espécies (Figura 6B). Quando comparados apenas os internódios novos, as duas primeiras espécies também apresentaram os maiores do que as duas outras. Porém, os internódios maduros de *S. barberi*, *S. officinarum* e *S. spontaneum* foram maiores do que o de *S. robustum*.

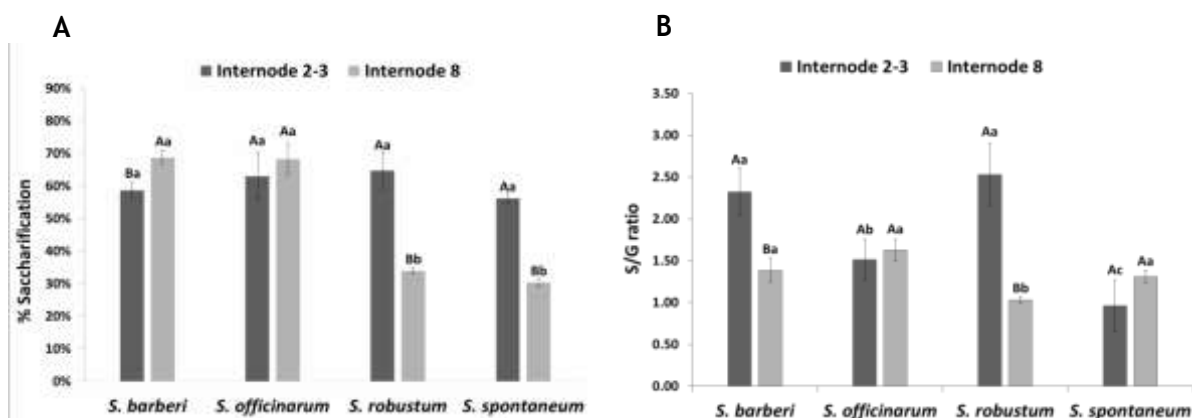


Figura 6. Porcentagem de sacarificação (A) e razão S/G (B) nos internódios de espécies de *Saccharum*. Letras maiúsculas diferentes denotam diferenças significantes entre internódios de diferente estágio de desenvolvimento dentro de uma mesma espécie. Letras minúsculas diferentes indicam diferenças significantes entre internódios de um mesmo estágio de desenvolvimento das diferentes espécies. As médias foram comparadas pelo teste de Tukey,  $P \leq 0,05$ . As barras verticais indicam o erro padrão das médias de quintuplicatas independentes.

## 5. 6. Perfil de oligômeros da lignina

A identificação dos monômeros e oligômeros de ligninas foi feita por comparação com dados de uma biblioteca “Do-It-Yourself” (Kiyota et al., 2012), usando os tempos de retenção, da razão  $m/z$  e padrão de fragmentação MS/MS. Foram encontradas estruturas de ligação pertencentes aos grupos  $\beta$ -aril éter (8-O-4), fenilcumaranas (8-5) e resinol (8-8) e foi possível identificar a unidade aromática S envolvida em cada umas destas estruturas de ligação (tabela 2). Nas espécies em estudo foram identificados 11 estruturas: um aldeído (S), um monolignol (S), quatro dímeros e cinco trímeros. Dois dos dímeros e dois dos trímeros apresentaram ligações 8-5, com os monômeros G que resultam em ligações mais recalcitrantes. As outras ligações presentes nos trímeros (8-0-4) são caracterizadas como as de mais fácil clivagem. Dois dos dímeros ( $m/z=357$ , G(8-5)G e  $m/z= 387$ , S(8-5)G; S(8-8)G) apresentaram estereoisomeria (tempo de retenção= 3.39, 4.08 e 3.42, 3.39 respectivamente). Esta característica foi também apresentada para dois dos cinco trímeros identificados neste

estudo ( $m/z = 583$ , G(8-O4)S(8-5)G e  $m/z = 643$ , S(8-O4)S(8-8)S, tempos de retenção: 3.57, 3.66 e 3.82, 4.2 respectivamente).

Tabela 2. Oligômeros precursores de ligninas, tempo de retenção e suas respectivas massa-carga ( $m/z$ ) obtidos por UPLC-MS/MS em internódios de espécies de *Saccharum*.

Unit	Struture	$m/z$	Retention Time (min)
Monomers	Sinapyl aldehyde	207	3.26
	Sinapyl alcohol	209	2.75
Dimers	G(8-5)G	357	3.39 e 4.08
	G(8-O-4)G	375	3,76
	S(8-5)G; S(8-8)G	387	3.42 e 3.9
	G(8-O-4)S	405	3,2
Trimers	G(8-O-4)G(8-5)G	553	3,7
	G(8-O4)S(8-5)G	583	3.57 e 3.66
	G(8-O-4)S(8-O-4)G	601	3,58
	S(8-O-4)S(8-O-4)G	631	3,65
	S(8-O4)S(8-8)S	643	3.82 e 4.2

Independentemente da idade dos internódios, *S. spontaneum* e *S. robustum* possuem a maior diversidade e frequência de oligômeros quando comparadas com *S. officinarum* e *S. Barberi* (Figura 7). O aldeído S ( $m/z = 207$ ) foi encontrado tanto para internódios jovens como para maduros de todas as espécies em estudo, enquanto que o monolignol S ( $m/z = 209$ ) foi encontrado só nos internódios jovens de *S. officinarum* e *S. barberi*, sendo com maior frequência para esta última espécie. Nas espécies *S. robustum* e *S. spontaneum*, os dímeros de lignina apresentam uma tendência de prevalecer nos internódios maduros, de forma oposta ao encontrado para *S. officinarum* e *S. barberi*, onde os dímeros se identificam com maior frequência para os internódios jovens. Os trímeros foram identificados preferencialmente nos internódios maduros das quatro espécies, sendo com notável frequência em *S. spontaneum* e *S. robustum*. Comparando todos os oligômeros identificados, o dímero  $m/z = 387$  [S(8-5)G; S(8-8)G] e o trímero  $m/z = 583$  [G(8-O4)S(8-5)G] foram as estruturas que apresentaram maior frequência, e de forma oposta, a menor frequência foi para o dímero  $m/z = 405$  [G(8-O-4)S]. O dímero  $m/z = 357$  [G(8-5)G] foi encontrado só nas espécies *S. robustum* e *S. spontaneum*. Não foram identificados os dímeros  $m/z = 357$  [G(8-5)G] e  $m/z =$

405 [G(8-O-4)S)] tanto no internódio jovem como no internódio maduro das espécies *S. officinarum* e *S. barberi*.

A

Species	Age	m/z										
		357	207	209	375	387	405	553	583	601	631	643
<i>S. barberi</i>	Internode 2+3											
<i>S. officinarum</i>	Internode 2+3											
<i>S. robustum</i>	Internode 2+3											
<i>S. spontaneum</i>	Internode 2+3											

B

Species	Age	m/z										
		357	207	209	375	387	405	553	583	601	631	643
<i>S. barberi</i>	Internode 8											
<i>S. officinarum</i>	Internode 8											
<i>S. robustum</i>	Internode 8											
<i>S. spontaneum</i>	Internode 8											

Frequency of detection	0	x1	x2	x3	x4	x5

Figura 7. Distribuição de oligômeros precursores de ligninas e suas respectivas massa-carga ( $m/z$ ) nos internódios de diferentes idades em espécies de *Saccharum*. A frequência de cada estrutura é representada no diagrama por diferentes intensidades da cor verde, indo de não encontrada (0 - branco) até encontrada em todas as cinco amostras analisada (x5 - verde escuro intenso).

### 5.7. Fenóis totais

Os maiores teores de fenóis foram encontrados nos internódios mais novos, sendo que os maiores valores foram os das espécies *S. robustum* e *S. spontaneum*. *S. officinarum* apresentou o menor teor de fenóis nos internódios novos. As quatro espécies não diferiram quanto ao teor nos internódios maduros.

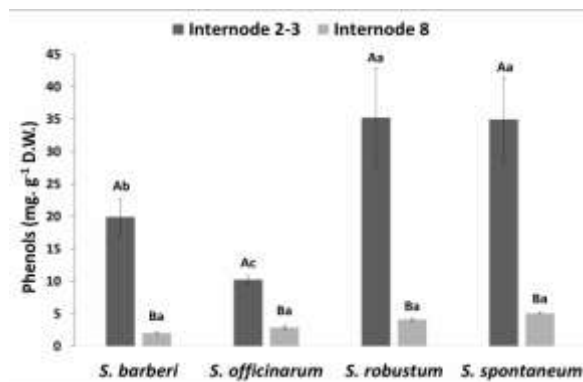


Figura 8. Conteúdo de Fenóis totais em internódios de espécies de *Saccharum*. Letras maiúsculas diferentes denotam diferenças significantes entre internódios de diferente estágio de desenvolvimento dentro de uma mesma espécie. Letras minúsculas diferentes indicam diferenças significantes entre internódios de um mesmo estágio de desenvolvimento das diferentes espécies. As médias foram comparadas pelo teste de Tukey,  $P \leq 0,05$ . As barras verticais indicam o erro padrão das médias de quintuplicatas independentes.

### 5.8. Composição de monossacarídeos, de lignina e acetilação da fração de xilano da parede celular

As Figuras 9A e 9B mostram a expansão do espectro 2D HSQC NMR ( $^1\text{H}$  (x-axis) /  $^{13}\text{C}$  (y-axis)) da região aromática de lignina e região anomérica, respectivamente de uma amostra da parede celular de internódio de uma das espécies. Picos proeminentes correspondentes a ligações polissacarídicas conhecidas estão indicados (Cheng et al. 2013; Chong et al., 2014).

As composições de *p*-hidroxicinamatos, grupos O-acetil substituintes em xilose e monossacarídeos são mostradas nas Figuras 9C, 9D e 9E. Não houve diferença significativa quanto a *p*-coumarato e ferulato (Figura 9C). Para a abundância relativa de grupos O-acetil substituintes de xilanos (Figura 9D), *S. officinarum* teve significativamente uma maior porcentagem de grupos substituintes 3-OAc em relação às outras espécies em estudo. Por outro lado, no que diz respeito ao grupo 2,3-OAc, houve diferenças significativas entre as espécies em estudo, sendo que a maior porcentagem foi encontrada em *S. officinarum* e *S. spontaneum*. Não se apresentam diferenças significativas entre as espécies em estudo, no que respeita à abundância relativa total dos grupos acetilados e o grupo substituinte 2-OAc. *S. spontaneum* e *S. robustum* apresentaram significativamente o maior conteúdo de glicose, com relação às espécies *S. officinarum* e *S. barberi*. De forma oposta ao encontrado para a glicose, o porcentagem de xilose foi significativamente maior nas espécies *S. officinarum* e *S. barberi*. *S. officinarum* apresenta significativamente a maior abundância de manose quando comparados com as outras espécies em estudo. Para o caso da arabinose, *S. barberi* foi a



espécie que apresenta o maior porcentagem deste monossacarídeo. *S. spontaneum*, *S. robustum* e *S. officinarum* não apresentam diferenças significativas com relação ao monossacarídeo arabinose.

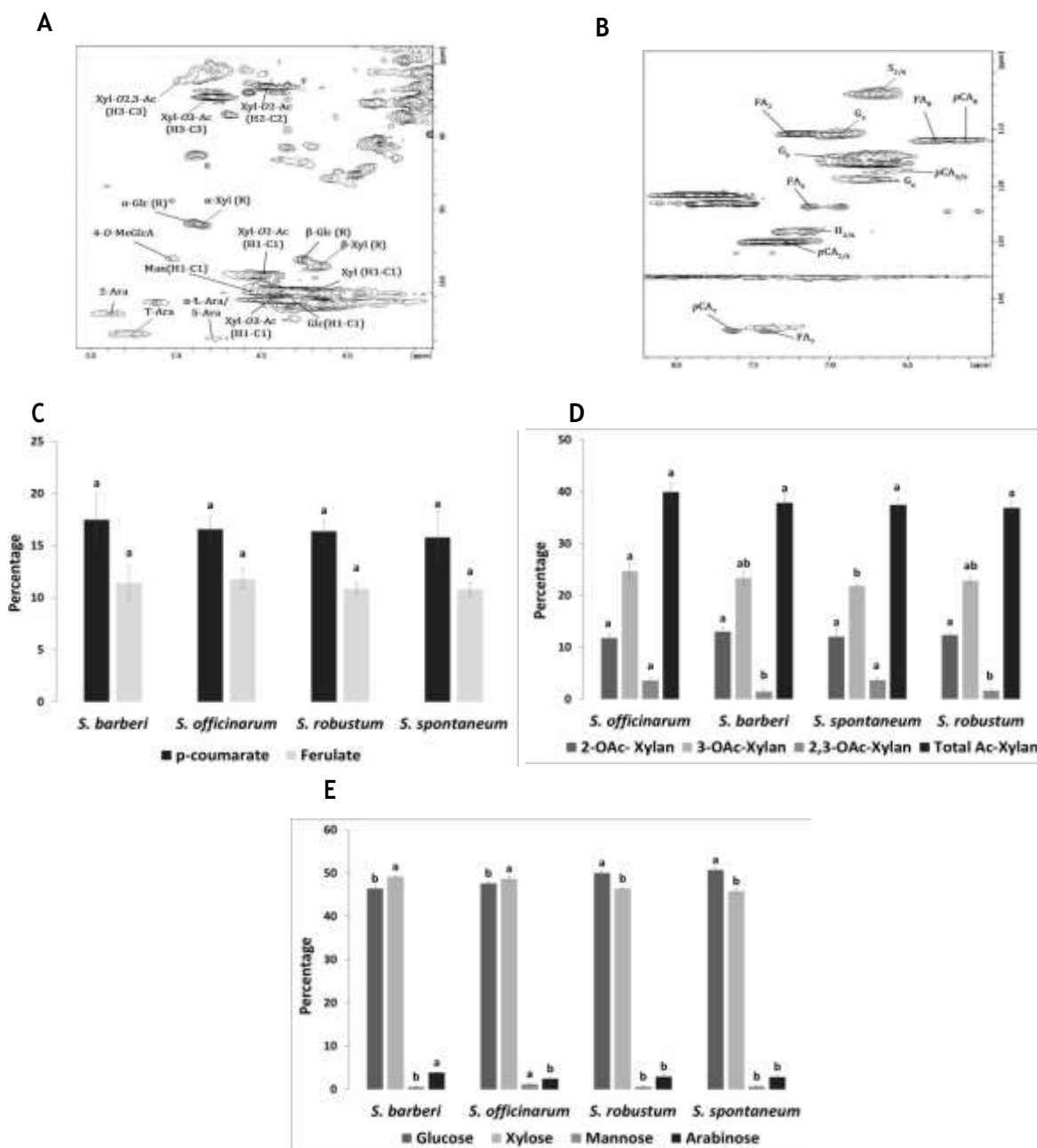


Figura 9. Espectros 2D HSQC NMR ( $^1\text{H}$  (x-axis) /  $^{13}\text{C}$  (y-axis)) da região anomérica (A) e região aromática da lignina (B) em vástagos de espécies de *Saccharum*. Porcentagem de *p*-hidroxicinamatos (C), grupos O-acetil substituintes (D) e monossacarídeos (E) em vástagos de espécies de *Saccharum*. As letras minúsculas diferentes indicam diferenças significativas entre os vástagos das espécies de *Saccharum* para um tipo de *p*-hidroxicinamato, grupo O-acetil substituinte ou monossacarídeo. As médias foram comparadas pelo teste de Tukey,  $P \leq 0,05$ . As barras verticais indicam o erro padrão das médias de triplicatas independentes.

### 5.9. Análises histoquímicas - lignina, celulose e amido

A identificação de lignina total foi feita com Fluoroglucinol-HCl e a presença de unidade S com o reagente de Maüle. O reagente de Maüle confere cor avermelhada à lignina siringil (S) e amarela à lignina guaiacil (G). Foram feitos cortes a mão livre dos internódios 2 (novo), 5 (intermediário) e 7 (maduro). Foram identificadas duas regiões bem definidas: córtex (constituído predominantemente por feixes vasculares, fibras do xilema e esclerenquimáticas) e medula (parênquima). Observou-se que com o aumento da maturidade houve aumento de lignina G (Figura 10) e lignina total (Figura 11), preferencialmente no córtex. Nos internódios jovens (Figuras 10-2<sup>nd</sup> e 11-2<sup>nd</sup>), a lignificação está restrita ao protoxilema tanto para a região do córtex e medula, sendo a lignina S a subunidade que se encontra preferencialmente nestes internódios (Figura 10-2<sup>nd</sup>). Por outro lado é importante assinalar que a reação com fluroglucinol foi mais intensa em *S. spontaneum* e *S. robustum* (Figuras 11-2<sup>nd</sup>-E-F e G-H, respectivamente). A deposição de lignina total nos internódios intermediários (Figuras 11-5<sup>th</sup>) apresenta diferenças entre as espécies, indicando que para *S. officinarum* e *S. robustum* (Figuras 11-5<sup>th</sup>-A-B e G-H, respetivamente) ocorreu no protoxilema, metaxilema, epidermis e na primeira camada de fibras ao redor dos feixes vasculares corticais, sendo a lignina S encontrada preferencialmente nestes tipos celulares (Figuras 10-5<sup>th</sup>-A-B e G-H). Para *S. spontaneum* e *S. barberi* (Figuras 11-5<sup>th</sup>-E-F e C-D, respetivamente) além dos tipos celulares lignificados descritos nas espécies anteriormente mencionadas, a deposição deste polímero aromático ocorre em várias camadas de fibras do feixe vascular cortical, fibras do esclerênquima e iniciando-se nas células parenquimáticas da medula e parênquima cortical, sendo a lignificação mais intensa, indicando que nestes tipos celulares começa se acumular precocemente a lignina G (Figuras 10-5<sup>th</sup>-E-F e C-D, respetivamente). *S. spontaneum* e *S. robustum* (Figuras 11-5<sup>th</sup>-E-F e G-H, respetivamente) apresentaram lignificação nas células do parênquima clorofiliano com relação a *S. officinarum* e *S. barberi* (Figuras 11-5<sup>th</sup>-A-B e C-D, respetivamente). No internódio maduro (7<sup>th</sup>) se observou que a lignificação ocorre em todos os tipos celulares, apresentando a tendência de acumular lignina G nas células do feixe vascular cortical e fibras do esclerênquima (Figuras 10 e 11-7<sup>th</sup>). Em *S. robustum* e *S. spontaneum* (Figuras 11-7<sup>th</sup>-G-H e E-F, respetivamente) os tipos celulares apresentam paredes mais espessas e lignificadas do que o observado em *S. barberi* e *S. officinarum* (Figuras 11-7<sup>th</sup>-C-D e A-B, respetivamente), indicando também que

os feixes vasculares da medula começam a apresentar deposição de lignina G (Figuras 10-7<sup>th</sup>-G-H e E-F).

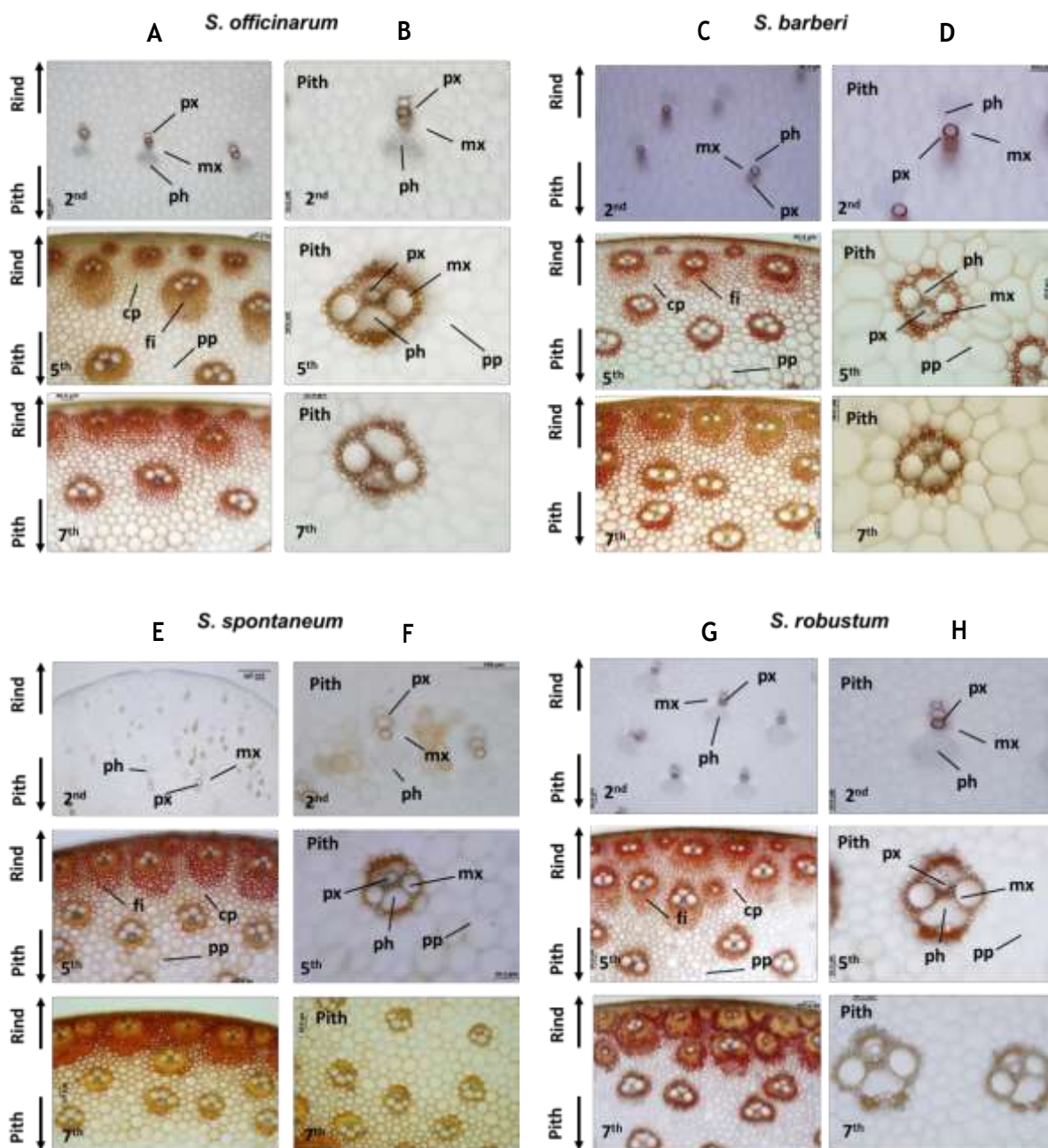


Figura 10. Secções transversais de diferentes regiões e internódios em espécies de *Saccharum* submetidas à reação de Maule para detecção de lignina S e G. 2<sup>nd</sup>: Internódio jovem, 5<sup>th</sup>: Internódio intermediário, 7<sup>th</sup>: Internódio maduro. A-C-E-G: Zona cortical (Rind); B-D-F-H: Zona medular (Pith). ph: Floema; px: protoxilema; mx: metaxilema; fi: fibras; cp: parênquima cortical; pp: parênquima medular. Barra = 50 μm.

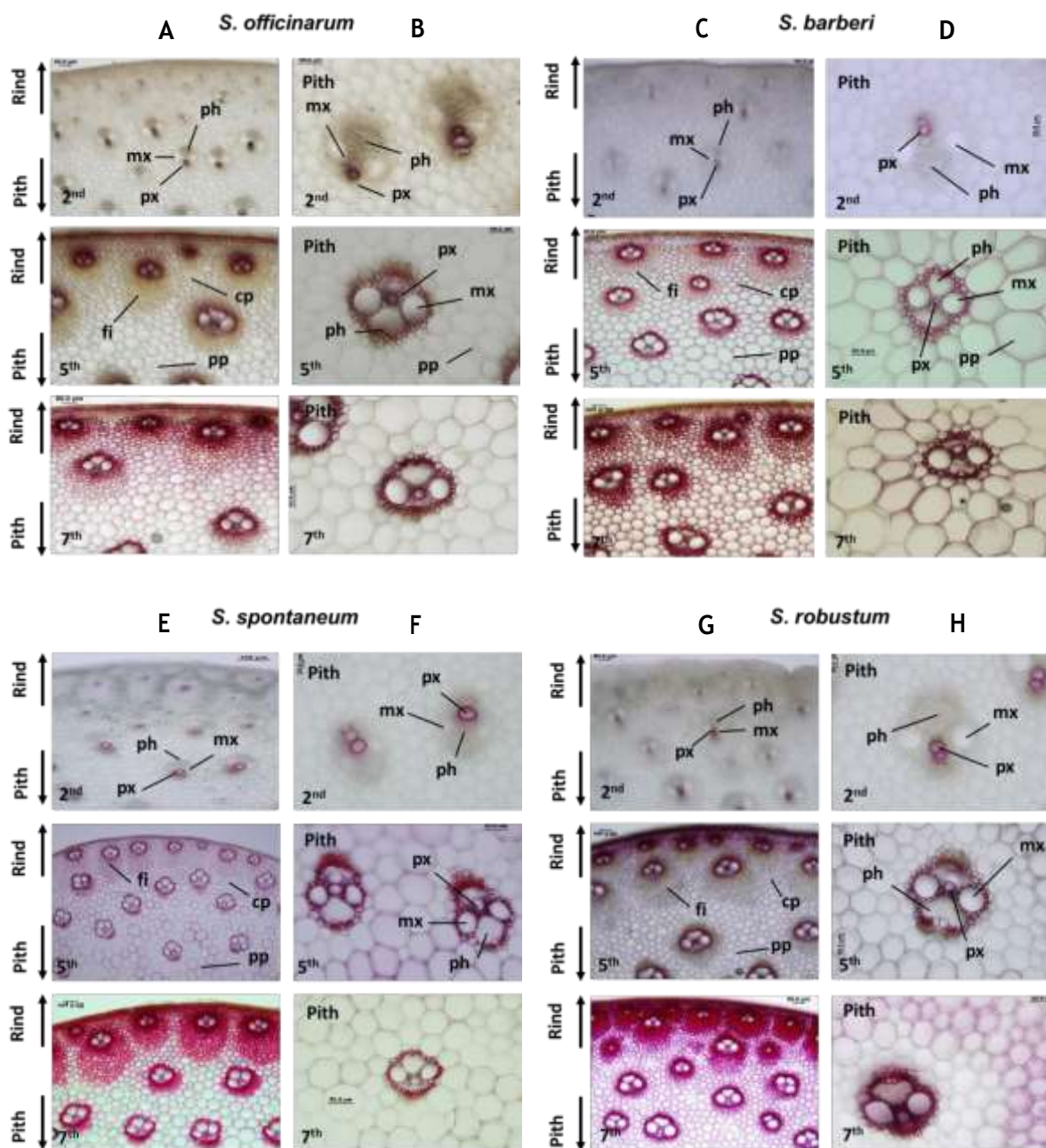


Figura 11. Secções transversais de diferentes regiões e internódios em espécies de *Saccharum* coradas com Fluoroglucinol para detecção de lignina total. 2<sup>nd</sup>: Internódio jovem, 5<sup>th</sup>: Internódio intermediário, 7<sup>th</sup>: Internódio maduro. A-C-E-G: Zona cortical (Rind); B-D-F-H: Zona medular (Pith). ph: Floema; px: protoxilema; mx: metaxilema; fi: fibras; cp: parênquima cortical; pp: parênquima medular. Barra = 50 μm.

Com relação à identificação de celulose, se observa que existem diferenças qualitativas entre os internódios jovens (2<sup>nd</sup>) e maduros (7<sup>th</sup>), indicando que a reação com calcoflúor foi mais intensa para a fase jovem (Figura 12-2<sup>nd</sup>). Por outro lado, é importante indicar que para os internódios jovens de todas as espécies, a identificação deste polímero foi mais evidente no córtex (Figuras 12-2<sup>nd</sup>-A-C-E-G), e de forma oposta, a medula foi a região que apresentou maior intensidade na visualização da celulose nos internódios maduros (Figuras 12-7<sup>th</sup>-B-D-F-H). No córtex e a medula do internódio jovem, de *S. robustum* (Figuras 12-2<sup>nd</sup>-G-H, respetivamente), todos os tipos celulares foram positivos para o teste de detecção de celulose, mas para o caso de *S. officinarum*, *S. barberi* e *S. spontaneum* a identificação da celulose foi negativa tanto para o protoxilema (córtex e medula) e a primeira camada de fibras do feixe vascular cortical. A ausência de visualização da celulose neste último tipo celular foi mais acentuada em *S. spontaneum*. Nos internódios maduros de *S. officinarum* e *S. barberi* (Figura 12-7<sup>th</sup>-A e C, respetivamente), se observou que a identificação da celulose no córtex ficou limitada só nas células do parênquima clorofiliano, sub-epidermicas e do floema, enquanto que para *S. robustum* e *S. spontaneum* (Figuras 12-7<sup>th</sup>-G e E, respetivamente) foi só para o floema. Na medula dos internódios maduros, a celulose só foi visualiza no parênquima de armazenamento, sendo isto mais evidente para *S. officinarum* (Figura 12-7<sup>th</sup>-B). Deve-se notar porém que para *S. robustum* (Figura 12-7<sup>th</sup>-H) as fibras do esclerênquima próximo à fronteira da região medular também mostraram uma discreta visualização de celulose.



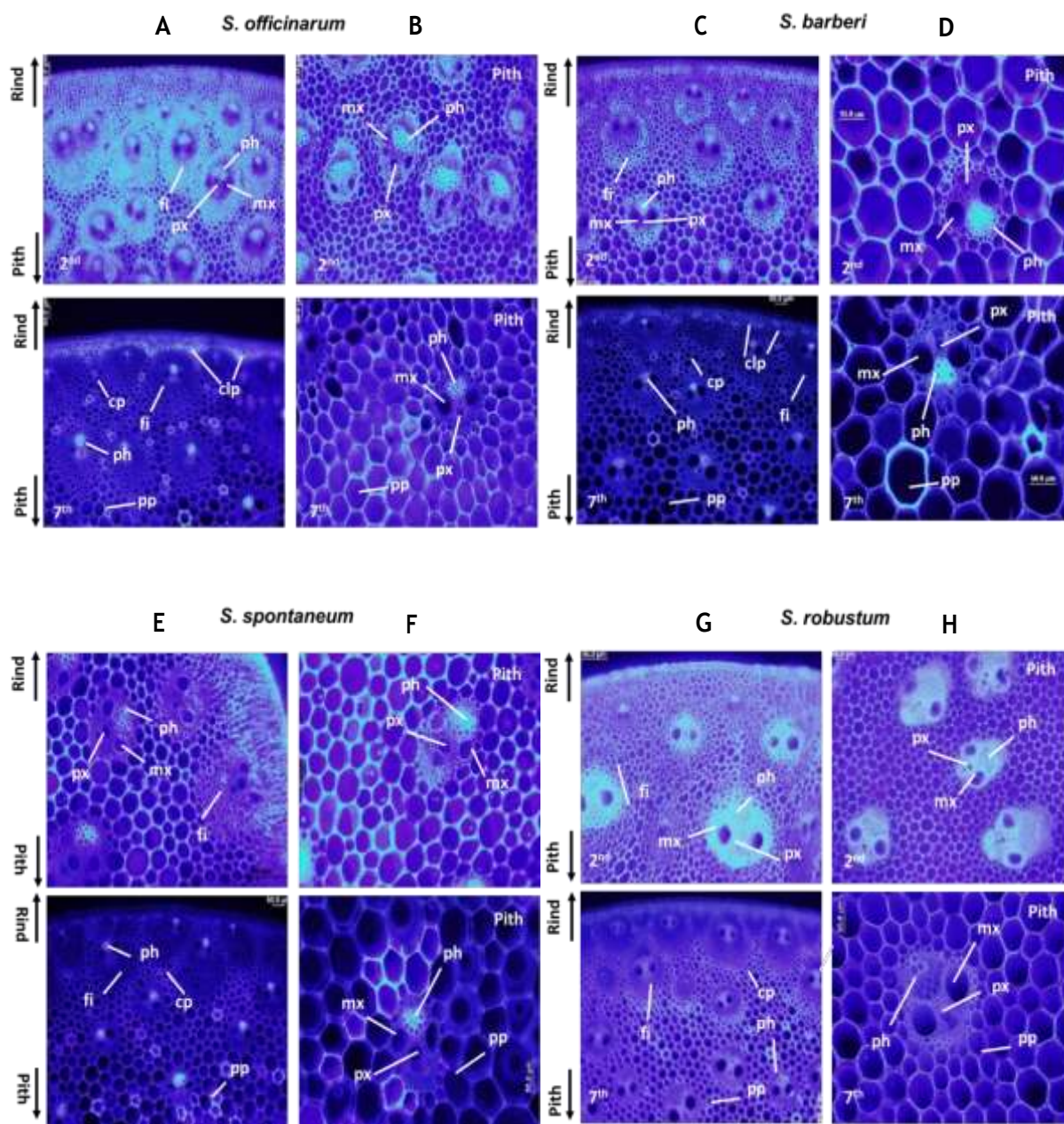


Figura 12. Secções transversais de diferentes regiões e internódios em espécies de *Saccharum* corada com Calcofluor-White para detecção de Celulose. 2<sup>nd</sup>: Internódio jovem e 7<sup>th</sup>: Internódio maduro. A-C-E-G: Zona cortical (Rind); B-D-F-H: Zona medular (Pith). ph: Floema; px: protoxilema; mx: metaxilema; fi: fibras; cp: parênquima cortical; pp: parênquima medular e clp: parênquima clorofiliano. Barra = 50 µm.

A coloração com lugol revelou diferenças na presença de grânulos de amido entre os internódios jovens (Figura 13-2<sup>nd</sup>), das espécies em estudo, indicando que a identificação deste polissacarídeo só foi positiva para *S. officinarum* e *S. robustum* (Figuras 13-2<sup>nd</sup>-A e D,

respetivamente), sendo que o acúmulo de amido foi discreto nestas espécies e se limitou aos tipos celulares pertencentes ao parênquima clorofiliano e ao parênquima cortical. Para os internódios maduros de *S. officinarum*, *S. robustum* e *S. barberi* (Figuras 13-7<sup>th</sup>-A-D e B) verificou-se uma discreta visualização de amido que ocorre no parênquima clorofiliano, enquanto que para *S. spontaneum* (Figura 13-7<sup>th</sup>-C) o acúmulo foi notável com relação às outras espécies descritas anteriormente, indicando que o amido se visualiza nas células do parênquima clorofiliano, cortical e medular.

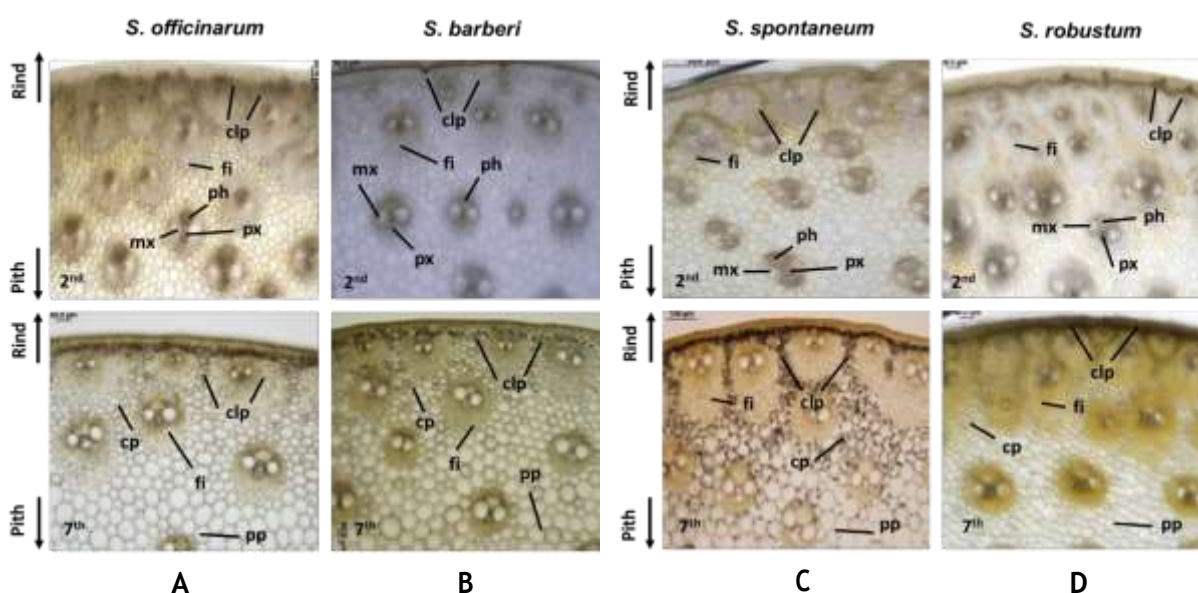


Figura 13. Secções transversais de diferentes regiões e internódios em espécies de *Saccharum* tratadas com lugol ( $I_2+KI$ ) para detecção de amido. 2<sup>nd</sup>: Internódio jovem e 7<sup>th</sup>: Internódio maduro. ph: Floema; px: protoxilema; mx: metaxilema; fi: fibras; cp: parênquima cortical; pp: parênquima medular e clp: parênquima clorofiliano. Barra = 50  $\mu m$ ; 2<sup>nd</sup>C= 200  $\mu m$  e 7<sup>th</sup>C=100  $\mu m$ .

### 5.10. Identificação de genes associados com a rota biossintética da lignina

Os genes da biossíntese de lignina identificados em híbridos comerciais de cana (Bottcher et al., 2013) foram usados como iscas em bancos genômicos de dados. Entre as sequências que retornaram, foram selecionadas CDs completos aquelas de sorgo, arroz, milho, trigo, azevém e *A. thaliana*. Estas sequências foram alinhadas para identificação de primers conservados, que foram então usados em reações de RT-PCR. Quando isto não foi possível, foram desenhados primers degenerados. As bandas amplificadas foram sequenciadas e sua identidade verificada por comparação com sequências dos bancos de dados. Um total de 13 unigenes foram identificados nestas espécies, os quais foram: 1 *C4H*, 2 *4CL*, 1 *HCT*, 1 *F5H*, 1

*C3H*, 2 *CCoAOMT*, 1 *CCR*, 1 *COMT* e 3 *CAD*. Estas sequências foram traduzidas para a correspondente sequência de aminoácidos e usadas em análises filogenéticas com o objetivo de identificar homólogos que possam ter papel na lignificação nas espécies em estudo e para definir a possível existência de isoformas dentro de cada família de genes.

### 5.10.1. Filogenia de C4H

C4H pertence ao grupo CYP73A da família do citocromo P450 e catalisa a conversão do ácido cinâmico para ácido *p*-coumárico (Meijer et al., 1993). Foi descrito só um gene C4H para *A. thaliana* (Mizutani et al., 1997), enquanto que para *Z. mays* (Guillaumie et al., 2007), *P. virgatum* (Shen et al., 2013) e cana de açúcar (Bottcher et al., 2013) foram identificados dois genes. *S. bicolor* e *O. sativa* apresentam três e quatro genes codificando C4H, respetivamente (Xu et al., 2009). Aqui foi identificado apenas um gene codificando para C4H, que (Figura S1). A análise filogenética mostrou a formação de três clados, sendo que o primeiro é conformado exclusivamente por monocotiledôneas, incluindo a sequência de C4H isolada das espécies em estudo e a SAS de cana de açúcar SCCCL4009H01.g (*ShC4H1*, Bottcher et al., 2013), encontrando-se estreitamente associados com as sequências de *Z. mays* C4H ACG34412.1, C4H NP 001149158.1, *S. bicolor* C4H AAK54447.1 e *P. virgatum* C4H1 AFY17064.1. O segundo clado inclui sequências tanto de dicotiledôneas como de monocotiledôneas. O terceiro clado foi formado somente pelas sequências de *Z. mays* C4H NP 001151365 e *Nicotiana tabacum* C4H AAK62344.1. De forma divergente aos outros clados pertencentes da árvore filogenética da família do gene C4H neste estudo, a SAS de cana de açúcar SCJFRZ2030D03 (*ShC4H2*, Bottcher et al., 2013) se encontra formando um grupo separado.

O alinhamento das sequências traduzidas de aminoácidos mostrou que, com exceção de apenas um aminoácido em *S. officinarum*, houve 100% de identidade entre as espécies em estudo e o gene *ShC4H1* identificado por Bottcher et al. (2013) (Figura S1A).

### 5.10.2. Filogenia de 4CL

4CL é uma enzima que catalisa a formação de ésteres de CoA dos ácidos cinâmicos: *p*-coumarico, caféico, ferúlico e 5-hidroxi-ferúlico e sinapico. 4CL desempenha um papel fundamental na canalização de precursores fenilpropanóides em diferentes vias, cada uma conduzindo a uma variedade de produtos finais funcionalmente distintos, incluindo



ácido benzóico, taninos condensados, flavonóides e cinamil álcoóis (Gross, 1985; Harding et al., 2002). Para *Z. mays* (Puigdomenech et al., 2001) e *P. virgatum* (Shen et al., 2013) foram identificados duas sequências de 4CL. Em cana de açúcar foram identificados 3 genes codificando para 4CL (Bottcher et al., 2013), enquanto que para *A. thaliana* e *S. bicolor*, foram identificados 13 e 15 genes, respectivamente (Xu et al., 2009). Neste estudo, foi identificado dois genes 4CL para a espécie *S. barberi* (denominados aqui: *S. barberi* 4CL A e 4CL B), enquanto que para as outras espécies só se identificou um gene, 4CL A (Figura S2). Da totalidade dos *reads* analisados na espécie *S. barberi* o gene 4CL A possui um 80% de abundância e 20% 4CL B 20%. Da árvore filogenética da família 4CL se verificou que é composta de 4 clados. O primeiro clado contém exclusivamente monocotiledôneas, incluindo as SASs de cana de açúcar (SCCCCL3002A03.g e SCVPLB1015F12.g) e as sequências 4 CL A das espécies em estudo. A sequência traduzida de 4CL A mostrou que, com exceção a uns poucos aminoácidos, foi praticamente idêntica à SAS SCCCCL3002A03.g (Figura S2A) identificada por Bottcher et al. (2013) como *Sh4CL1*. O terceiro clado está composto exclusivamente por dicotiledôneas. O quarto clado da árvore filogenética está composta por sequências de monocotiledôneas e dicotiledôneas, sendo que neste clado está inclusa a sequência de *S. barberi* 4CL B e a SAS de cana de açúcar SCMCRT2102F02.g (*Sh4CL3*, Bottcher et al., 2013).

### 5.10.3. Filogenia de HCT

Hidroxicinamoil-CoA.transferase é uma aciltransferase que catalisa a conversão do *p*-coumaroil-CoA e caffeoil-CoA nos seus correspondentes ésteres de shikimato e quinato (Hoffmann et al., 2003). Em *Arabidopsis* (Raes et al., 2003) e *Z. mays* possuem dois genes HCT (Guillaumie et al., 2007), enquanto que *S. bicolor* apresenta 4 (Xu et al., 2009). Em cana de açúcar foram identificados dois genes, sendo um deles HCT-like (Bottcher et al., 2013). Neste estudo encontramos nas quatro espécies apenas uma HCT, cujas sequências agruparam com a SAS SCCCCL4009E02.g, identificado por Bottcher et al. (2013) como *ShHCT1* (Figura S3). O alinhamento das sequências (Figura S3A) mostrou que, novamente, com pequenas alterações nas sequências de aminoácidos, os genes aqui identificados são praticamente idênticos a *ShHCT1*.

#### 5.10.4. Filogenia de C3H

C3H é uma enzima que pertence à família do citocromo P450 CYP98 e catalisa a conversão dos ésteres de shiquimato e quinato do ácido *p*-coumárico, dentro do seus correspondentes conjugados de ácido caféico (Schoch et al., 2001; Franke et al., 2002). No genoma de *A.thaliana* foram identificados três genes C3H (Raes et al., 2003), enquanto que para *Z. mays* (Guillaumie et al., 2007) e *O. sativa* (Xu et al., 2009) um só gene C3H foi encontrado. *S. bicolor* (Xu et al., 2009), cana de açúcar (Bottcher et al., 2013) e *P. virgatum* (Shen et al., 2013) apresentam dois genes C3H. Só um gene de C3H foi isolado nas espécies em estudo (Figura S4). A árvore filogenética construída apresentou três clados. O primeiro clado é formado exclusivamente por monocotiledôneas, incluindo a sequência C3H isolada das espécies em estudo, se encontrando, estreitamente relacionada ao gene *ShC3H1* (Bottcher et al., 2013; SAS SCVPCL6041E07.g) e C3H AAC39316.1 de *S. bicolor*. O segundo grupo da árvore foi formado só por sequências de dicotiledôneas. O terceiro grupo é divergente aos outros clados da árvore e está conformado por duas sequências de *A. thaliana*. Novamente, quando alinhadas as sequências traduzidas da C3H isolada nas quatro espécies apresentaram grande identidade com o gene *ShC3H1* (Bottcher et al., 2013) diferindo em poucos aminoácidos.

#### 5.10.5. Filogenia de CCoAOMT

CCoAOMT é uma enzima que catalisa a metilação do cafeoil-CoA em feruol-CoA (*in vivo* e *in vitro*) e 5-hidroxiferuol-CoA em sinapoil-CoA (*in vitro*), e conjuntamente com COMT, são as enzimas responsáveis pela metilação dos precursores dos monolignóis (Ye et al., 1994). *Z.mays* apresenta cinco genes codificando CCoAOMT (Guillaumie et al., 2007), enquanto *S.bicolor* e *A. thaliana* 7 genes (Xu et al., 2009). *P. virgatum* tem quatro genes (Shen et al., 2013), *O. sativa* 11 genes (Xu et al., 2009) e cana de açúcar 3 genes (Bottcher et al., 2013). Neste estudo foram identificados dois genes para as espécies de *Saccharum*, denominados aqui CCoAOMT A e CCoAOMT B (Figura S5). Da totalidade dos *reads* analisados na espécies estudadas as abundâncias de CCoAOMT A e CCoAOMT B nas espécies estudadas foram: 80% e 20% para *S. barberi*, 30% e 69% para *S. officinarum*, 26% e 73% para *S. robustum*, e 50% e 50% para *S. spontaneum*. Os dois genes apresentaram 60% de identidade (Figura S5A). Da análise filogenética obtida para o gene CCoAOMT pode-se formar quatro clados. O primeiro clado está composto exclusivamente por sequências CCoAOMT de monocotiledôneas, incluindo as sequências CCoAOMT A isoladas das

espécies em estudo, as SASs de cana SCJLRT2050C09, SCJFRZ2010H06.g e as sequências de *P. virgatum* CCoAOMT Fb01856.1 e Fb02068.1. As duas SAS de cana correspondem aos genes identificados por Bottcher et al. (2013) como *ShCCoAOMT2* e *ShCCoAOMT3*. A maior identidade em nível de aminoácidos das sequências das quatro espécies e as duas dos híbridos estudados por Bottcher et al. (2013) ocorreu com *ShCCoAOMT2* (Figura S5A). Com uma identidade de 88% ao nível de aminoácidos, a sequência de *P. virgatum* CCoAOMT Fb01856.1 se encontra estreitamente relacionada às sequências CCoAOMT A. O segundo clado, está composto somente por sequências CCoAOMT das monocotiledôneas representadas pelas espécies *P. virgatum*, *B. distachyon* e *O. sativa*. O terceiro clado é composto só por sequências CCoAOMT de dicotiledôneas. O quarto clado também é composto só por monocotiledôneas, como no caso do clado primeiro e segundo, mas nele estão as sequências isoladas de CCoAOMT B das espécies em estudo, a SAS SCCCLR1069B09.g (*ShCCoAOMT1*, Bottcher et al., 2013) de cana de açúcar e a sequência de *P. virgatum* JO4806.1, que se encontra associada com as duas últimas sequências mencionadas, apresentando uma identidade de 98% ao nível de aminoácidos.

#### 5.10.6. Filogenia de CCR

CCR, é uma enzima que catalisa a conversão dos ésteres de hidroxicinamoil-CoA (*p*-coumaroil-CoA, feruloil-CoA e sinapoil-CoA) nos seus correspondentes cinamil aldeídos, sendo a primeira atividade relacionada especificamente com a via biosintética dos monolignóis (Raes et al., 2003; Guillaumie et al., 2007). O genoma de *A. thaliana*, apresenta sete genes CCR (Raes et al., 2003), enquanto que *Z. mays* possui oito genes (Guillaumie et al., 2007), *P. virgatum* possui quatro (Shen et al., 2013) e dois em cana de açúcar (Bottcher et al., 2013). Só um gene de CCR foi identificado neste estudo (Figura S6) sendo que as sequências das quatro espécies estudadas apresentaram 100% de identidade. A análise filogenética mostrou a formação de quatro clados. O primeiro e terceiro clados estão formados exclusivamente por sequências de monocotiledôneas. No primeiro clado estão incluídas as CCR isoladas, a SAS SCCCRZ2C01A04.g (*ShCCR1*, Bottcher et al., 2013) de cana de açúcar e a CCR XP\_002445566.1. de *S. bicolor*. O alinhamento das sequências traduzidas de aminoácidos mostrou praticamente 100% de identidade entre as sequências aqui isoladas e *ShCCR1* (Figura S6A). O terceiro clado, contém só sequências CCRs de dicotiledôneas. De forma distante às sequências isoladas de CCR das espécies em estudo, no

quarto clado aparece a SAS de cana de açúcar SCCCL6024F07 (*ShCCR2*, Bottcher et al., 2013).

#### 5.10.7. Filogenia de F5H

F5H é uma enzima também conhecida como coniferaldeído 5-hidroxilase e é uma citocromo P-450 monooxigenase (CYP84) que é requerida para a produção de monômeros de siringil, sendo a enzima é responsável pela 5-hidroxilação do coniferaldeído e/ou coniferil álcool (Humphreys et al., 1999; Li et al., 2000; Humphreys and Chapple, 2002). O genoma de *A. thaliana* e *Z. mays* possuem dois genes F5H (Raes et al., 2003; Guillaumie et al., 2007); *Oryza sativa* possui 3 genes (Xu et al., 2009), enquanto um só gene de F5H foi encontrado em *P. virgatum* (Shen et al., 2013) e cana de açúcar (Bottcher et al., 2013). Neste estudo foi isolado um só gene de F5H (Figura S7), que exceto um aminoácido em *S. barberi* (Figura S7A), a sequências traduzidas praticamente foram iguais a *ShF5H1* (SAS SCJLRT1022E04.g, Bottcher et al., 2013). A análise filogenética mostrou a formação de 3 clados, indicando uma clara separação entre F5H de dicotiledôneas (clado II) e de monocotiledôneas (clado I e III). As sequências de *P. virgatum* (AFH89637 e BAO37287) ficaram colocadas exclusivamente no clado III da árvore filogenética. O primeiro clado contém a sequência F5H isolada das espécies em estudo, o SAS codificando F5H em cana de açúcar (SCJLRT1022E04.g) e relacionada fortemente a estas sequências também está inclusa a sequência de *S. bicolor* XP 002464383.1.

#### 5.10.8. Filogenia de COMT

COMT (E.C. 2.1.1.68) é uma enzima bifuncional que catalisa predominantemente a metilação do 5-hidroxiconiferaldeído e/ou 5-hidroxiconiferil álcool em sinapaldeído e/ou sinapil álcool, respectivamente (Osakabe et al., 1999; Li et al., 2008). O genoma de *Brachypodium distachum* possui 4 genes COMT (Wu et al., 2013). Em *Panicum virgatum* (Shen et al., 2013), *A. thaliana* (Raes et al., 2003), *Z. mays* (Guillaumie et al., 2007) e cana de açúcar (Bottcher et al., 2013), foi identificada somente uma sequência da COMT. Nas espécies de *Saccharum*, se isolou um só gene de COMT (Figura S8). A análise das sequências de COMT indicou que existe 100% de identidade entre distintas espécies em estudo (Figuras S8 e S8A). A árvore filogenética apresentou 4 clados, que evidencia a separação entre dicotiledôneas (clado IV) e monocotiledôneas (clados: I, II e III). O grupo III, foi formado só com COMTs de *B. distachyon*. O primeiro clado da árvore filogenética contém a COMT

isolada das espécies em estudo, a COMT de cana de açúcar (SCRFLR1012F12.g, *ShCOMT1*, Bottcher et al., 2013) e, estreitamente relacionados pelos valores de bootstrap, as sequências de ortólogos de *S. bicolor* AA043609.1, *Miscanthus sinensis* ADI24332.1 e *Z. Mays* Q06509.

#### 5.10.9. Filogenia de CAD

CAD é uma enzima que catalisa a redução de *p*-hidroxicinamaldeídos dentro dos seus correspondentes álcoois e é a enzima responsável pelo último passo na rota biosintética dos monolignóis (Raes et al., 2003; Guillaumie et al., 2007). *Z. mays* apresenta 6 genes CAD (Guillaumie et al., 2007), *A. thaliana* 9 genes CAD (Raes et al., 2003), *B. distachyon* 7 genes (Trabucco et al., 2013), *S. bicolor* 14 genes CAD (Xu et al., 2009), *P. virgatum* 2 genes CAD (Shen et al., 2013) e cana de açúcar 4 genes CAD (Bottcher et al., 2013). Neste estudo, foram identificados três genes, sendo os três foram encontrados em *S. robustum* (CAD A, CAD B e CAD C) e dois (CAD A e CAD B) nas outras espécies (Figura S9). Da totalidade dos *reads* analisados na espécies estudadas, em *S. barberi* o gene CAD A e CAD B apresentaram abundâncias de 86% e 13%, em *S. officinarum* as abundâncias foram de 73% e 26%, em *S. spontaneum* 80% e. Em *S. robustum*, para o qual foram encontradas três CADs as abundância foram de 46% para CAD A, 26 % para CAD B e 26% para CAD C. Na comparação entre as três equências foram encontradas 48% de identidade entre CAD A e CAD B, 61% entre CAD B e CAD C e 47% entre CAD A e CAD C. O posicionamento dessas CADs na árvore filogenética (Figura S9) e sua comparação das sequências traduzidas e aminoácidos dessas CADs (Figura S9A) mostrou que as formas A, B e C correspondem aos genes *ShCAD8*, *ShCAD2* e *ShCAD7*, respectivamente, identificados por Bottcher et al. (2013). Da análise filogenética da família do gene CAD houve formação de três clados, sendo o primeiro composto exclusivamente por monocotiledôneas, incluindo as sequências do gene CAD A das espécies em estudo, a SAS de cana de açúcar SCEQLR1029E05.g (*ShCAD8*) e a sequência de *S. bicolor* CAD XP 002462350.1, que está relacionada. Também no primeiro clado se encontrou o gene CAD C de *S. robustum*, associado à SAS de cana de açúcar SCCCLB1001F10.g (*ShCAD7*). O segundo clado, contem sequências de monocotiledôneas e dicotiledôneas, incluindo as sequências CAD B das espécies em estudo e a SAS de cana de açúcar SCEPRZ1011A02 (*ShCAD2*, Bottcher et al., 2013). Relacionadas ao gene CAD B das espécies em estudo foram encontradas as sequências dos ortólogos de *Z. mays* CAD 2 CAA06687.1 e *P. virgatum* CAD 2 ADO01602.1. O terceiro clado incluiu a SAS de cana de açúcar SCACHR1038E08.g (*ShCAD6*, Bottcher et al., 2013) associada estreitamente à

sequência de *S. bicolor* CAD XP 002446076.1. Neste último grupo, as únicas sequências de dicotiledôneas encontradas foram as de *A. thaliana* CAD9 NP 195643.1 e CAD AAP68279.1

## 6. DISCUSSÃO

O aumento do consumo mundial de recursos finitos de combustíveis fósseis e as consequências climáticas negativas do seu uso estão atualmente dirigindo a busca de fontes alternativas de energia: a biomassa vegetal. Cana de açúcar, é uma das plantas mais eficientes em coletar e converter a luz solar em energia química, apresentando uma alta capacidade de acumular biomassa. Recentemente, com o advento do desenvolvimento da tecnologia de E2G, o bagaço de cana de açúcar, material obtido após a extração do xarope rico em sacarose, tem recebido atenção como matéria prima para a produção desse biocombustível. O seu uso, porém, apresenta grandes desafios, desde que a biomassa de cana apresenta um alto conteúdo de lignina (~23%), que limita o processo de sacarificação pois aumenta a recalcitrância da parede celular e inibe a fermentação por liberar compostos tóxicos aos microrganismos fermentadores. Portanto, o estudo e entendimento dos fatores que governam a recalcitrância da parede celular em plantas que apresentam características de bioenergia poderia ser de relevante importância para o desenvolvimento de genótipos de variedades híbridas de cana com uma qualidade de biomassa para o seu uso em programas de bioenergia, de tal forma que a produção de E2G possa ser maximizada.

A finalidade desta pesquisa foi estudar a biossíntese de lignina em quatro espécies de *Saccharum*, a saber, *S. spontaneum*, *S. officinarum*, *S. robustum* e *S. barberi*, através da identificação de genes da via biossintética dos monolignóis e a caracterização dos componentes da parede celular ao nível histoquímico e bioquímico.

A eficiência de recuperar açúcares fermentáveis, para a produção de biocombustíveis a partir da biomassa vegetal está determinado grandemente pela composição de carboidratos. A biomassa vegetal compreende carboidratos não estruturais (açúcares solúveis e amido) e componentes estruturais da parede celular (celulose, hemicelulose e pectinas) (Arai-Sanoh et al., 2011). Portanto, compreender a diversidade de carboidratos estruturais e não estruturais, e a sua caracterização, em relação à variabilidade entre espécies, representa uma importante oportunidade para o desenvolvimento de paredes celulares, apresentando características ideais para a produção de etanol.

Os açúcares possuem um papel central no metabolismo das plantas, sendo que todo o carbono fixado pela fotossíntese passa por açúcares ou seus derivados (Wang et al., 2013). Algumas espécies, tais como cana de açúcar, com adaptações anatômicas-bioquímicas

especiais e fonte-drenos únicos, tem capacidade de armazenar no seus caules açúcares solúveis (principalmente sacarose) até 18% da massa fresca (Inman-Bamber et al., 2011). De forma oposta, os caules da maioria das outras espécies armazenam carbono como polissacarídeos insolúveis, tais como amido, celulose e hemicelulose, com uma baixa concentração de sacarose (Wang et al., 2013). Sacarose e os seus produtos hidrolisados, os açúcares redutores (glicose e frutose), representam a maior fração dos açúcares solúveis em cana de açúcar, tendo suas proporções alteradas nos diferentes estágios de maturação dos internódios, sendo negativamente correlacionados um com o outro (McCormick et al., 2006; Pan et al., 2009; Papini-Terzi et al., 2009). Em concordância com esta última afirmação, podemos indicar que nas espécies em estudo existiu uma correlação negativa entre estes tipos de açúcares, desde que a sacarose acumulou mais acentuadamente nos internódios maduros e os açúcares redutores preferencialmente nos internódios jovens (Figuras 4B e 4C). Por outro lado, dois grupos contrastantes de espécies foram observados com relação ao conteúdo de açúcares solúveis totais quando se comparados os internódios maduros. *S. officinarum* e *S. barberi* foram as espécies que apresentaram o maior conteúdo destes carboidratos (42% e 50%, respetivamente) quando comparadas com *S. spontaneum* e *S. robustum* (14% e 18%, respetivamente) (Figura 4A). Estas tendências observadas neste estudo são similares aos reportados por Bull and Glasziou, (1963), que trabalharam com colmos maduros de clones de *S. officinarum* e *S. sinensis*, que acumularam aproximadamente 50% (peso seco) de açúcares, enquanto que as espécies de *S. robustum* e *S. spontaneum* apresentaram valores de açúcares < 20%. Os baixos valores de açúcares solúveis encontrados para *S. spontaneum* e *S. robustum* em nosso estudo poderiam ser explicados pela alta demanda de carbono para a biossíntese da parede celular, através do rápido consumo de açúcares redutores na biossíntese de polissacarídeos estruturais complexos (Waclawovsky et al., 2010), não permitindo o acúmulo de sacarose em altas concentrações nos vacúolos das células parenquimáticas do colmo (Sato, 2012). Isto de certa forma encontra apoio no fato de que *S. robustum* e *S. spontaneum* apresentaram 45% e 48% de celulose nos seus internódios maduros, respetivamente, quando comparadas com *S. barberi* e *S. officinarum* (25% e 20%, respetivamente, Figura 3A).

Amido, é um polissacarídeo não estrutural produzido em plantas de cana de açúcar como substância de reserva (Imrie and Tilbury, 1972; Bewley and Black, 1994), podendo estar presente em colmos e em folhas, sendo mais abundante nestas últimas (Eggleston et al., 2007). Caules de cana de açúcar apresentam variáveis quantidades de amido, que dependem de vários fatores: genótipo, idade fisiológica, tecido ou órgão, área do caule e



meio ambiente. Por outro lado, no que diz respeito à idade fisiológica, conforme aumenta a maturidade o conteúdo de amido diminui (Zhou et al., 2008). O amido é considerado uma impureza em cana de açúcar, pois afeta adversamente o processamento, a qualidade e quantidade da açúcar destinada para a produção de bioetanol. Portanto, o conhecimento sobre o conteúdo deste polissacarídeo nas espécies de *Saccharum* é de grande importância para a seleção de parentais, que permitam desenvolver variedades híbridas com baixo teor de amido. Neste estudo se observa que o conteúdo de amido foi de maneira decrescente como segue: *S. spontaneum* > *S. officinarum* > *S. barberi* > *S. robustum* (Figura 4D), indicando que *S. spontaneum* apresentou de maneira notável o maior valor de amido quando comparada com as outras espécies de *Saccharum*. Nosso resultados está em concordância com dados publicados anteriormente (Dutt and Narasimhan, 1951; Zhou et al., 2007), onde foram analisados 49 e 215 clones, respectivamente, relacionados às espécies de *Saccharum* e outras como *Erianthus* e *Miscanthus*, e que *S. robustum* foi a espécie que menos acumulou amido (traços) enquanto que *S. spontaneum* apresentou o mais alto conteúdo. Por outro lado, com relação à idade do internódio, se pode observar que as espécies *S. officinarum*, *S. robustum* e *S. barberi*, apresentam uma preferência de acúmulo de amido nos internódios jovens. Contrariamente a este perfil de acúmulo de amido, *S. spontaneum* acumula notavelmente amido nos internódios maduros quando se compara com as outras espécies em estudo. O encontrado neste estudo para *S. officinarum*, *S. robustum* e *S. barberi* está corroborado com o apresentado por Ramakrishnan (1958), onde se verificou que todos os internódios, com exceção do internódio superior, só apresentava sacarose, glicose e frutose, mas não amido, para todos os estádios de desenvolvimento.

Nosso resultados com relação à *S. officinarum*, *S. robustum* e *S. barberi* nos levam a pensar que as rotas metabólicas sintetizando sacarose poderiam ser predominante em todos os internódios, mas nos internódios jovens a síntese de amido poderia ser uma característica especial. O acúmulo notável de amido nos internódios maduros de *S. spontaneum* estaria relacionado com a capacidade reconhecida de esta espécie para o perfilhamento e como uma estratégia para enfrentar eventos periódicos estresse bióticos e abióticos (Slewinski, 2012).

Os açúcares neutros são os principais blocos estruturais dos componentes da parede celular vegetal: celulose, hemicelulose e pectina. Estes carboidratos podem ser colocados dentro de dois grupos baseados segundo na abundância: (1) Alta: glicose (Glc),

xilose (Xyl), arabinose (Ara) e galactose (Gal) e (2) Baixa: ramnose (Rha), fucose (Fuc) e manose (Man). A relativa proporção dos açúcares neutros varia entre os distintos tipos de plantas e no desenvolvimento (Carpita, 1996; Fincher, 2009) e significativamente entre espécies de gramíneas forrageiras (Hatfield et al., 2009). A partir do análise por espectroscopia 2D HSQC NMR (tabela 2), podemos indicar que o conteúdo dos açúcares neutros das espécies em estudo foi contrastante com os outros reportado para outras gramíneas na literatura (Hatfield et al., 2009; Cheng et al., 2013). O porcentagem de xilose encontrado nos caules das espécies de *Saccharum* (45.8%-49.15 %) foi alto em comparação com a biomassa *Miscanthus* (~31%) e Cordgrass (~25 %). De forma oposta, a proporção de glicose (46.4%-50.71%), foi menor comparado com outras gramíneas (*Miscanthus* ~62%; Cordgrass ~66%). O conteúdo de manose (0,56%-1,23%) é também baixo quando comparado com outras gramíneas, mas o conteúdo de arabinose é alto (2,44%-3,88%). Por outro lado, quando comparando os caules das as espécies em estudo, se apresenta uma relação inversa entre o porcentagem de glicose e xilose. *S. spontaneum* e *S. robustum* apresentaram o maior conteúdo de glicose, enquanto que em *S. officinarum* e *S. barberi* o conteúdo de xilose foi alto. Isto pode ser explicado a partir de que uma baixa biossíntese de glicose frequentemente, é compensada com xilose e vice-versa, desde que estes monossacarídeos são os componentes estruturais da celulose e hemicelulose-pectina, que são drenos que competem por carbono (Bekker, 2007; Rancour et al., 2012).

Arabinose foi o terceiro monossacarídeo mais abundante nas espécies de *Saccharum*. *S. barberi* apresentou um porcentagem de arabinose significativamente mais alta do que as outras espécies em estudo. O conteúdo de arabinose está relacionado com a estrutura dos xilanos de gramíneas, desde que os xilanos apresentam uma grande variabilidade no que diz respeito ao seu grau de substituição com arabinose, determinado pela razão arabinose/xilose (Scobbie et al., 1993; Morrison et al., 1994). Possivelmente o maior conteúdo de arabinose observado em *S. barberi* se deva ao fato de que a estrutura dos xilanos nesta espécie apresenta mais sítios de substituição para a arabinose, corroborado pela sua razão arabinose/xilose (0.080), maior do que as outras espécies em estudo.

Os principais componentes estruturais da lignocelulose em monocotiledôneas incluem aos polissacarídeos celulose, hemicelulose e pectina e o polímero aromático lignina (Carpita, 1996), e está documentado que a proporção e o conteúdo destes componentes estruturais varia entre distintos grupos de espécies e ao longo do desenvolvimento (Carpita, 1996; Fincher, 2009). Comparando nossos resultados com os internódios maduros

verificamos que na espécie *S. officinarum* foram encontrados os valores mais baixos de celulose e hemicelulose (20,32% e 8,73%, respectivamente), enquanto que *S. spontaneum* os mais altos (46,15% e 17,18 %, respectivamente) (Figuras 3A e 3B). Os valores aqui encontrados de celulose nas espécies de *Saccharum* se encontram dentro do faixa observada para outras espécies de gramíneas relacionadas: *Miscanthus* (41.9%) (Magid et al., 2004), *Panicum* (46.1%) (Lynd et al., 1999) e cana de açúcar (39%) (Masarin et al., 2011). Porém, quando comparando o teor de hemicelulose com estas mesmas espécies, os valores encontrados aqui são mais baixos, o que poderia ser explicado pelo fato de terem sido materiais pré-tratados (pré-tratamento ácido ou alcalino) para posterior hidrólise enzimática, ou também porque nosso estudo se restringiu à análise do internódio. Por outro lado,, discutindo a idade do internódio, as espécies de *Saccharum* apresentam um perfil de acúmulo de celulose, variável, desde que em *S. officinarum* e *S. barberi* os internódios jovens apresentaram o maior conteúdo, enquanto que para *S. spontaneum* e *S. robustum* a deposição foi maior para os internódios maduros. Este último resultado está em concordância com dados obtidos para *Panicum virgatum* (Butkutė et al. 2013).

No que respeita à hemicelulose e pectina, todas as espécies em estudo apresentaram uma tendência geral de acúmulo nos internódios jovens (Figuras 3B e 3C). A diminuição de celulose e hemicelulose, em relação ao aumento da maturidade dos internódios em *S. officinarum*, coincide com o aumento na concentração de sacarose nesta espécie (Figura 4B), o que possivelmente é uma indicação de que o carbono que originalmente devia ser distribuído para a produção de polissacáridos estruturais foi utilizado para síntese de sacarose. Tal situação também foi observada também por Lingle and Thomson (2012), que caracterizou internódios de distintas idades em cinco genótipos de cana de açúcar, durante duas safras consecutivas. O fato de que existe uma correlação negativa no acúmulo de hemicelulose-pectina e celulose entre os internódios novos e maduros pode ser explicado a partir de que um balanço entre a biossíntese desses compostos durante o desenvolvimento da cana de açúcar (Sato et al., 2001).

Lignina é um polímero aromático tridimensional complexo que compromete uma substancial proporção (20%), das paredes celulares vegetal de gramíneas (Butkutė et al. 2013), atuando como uma união intermolecular e intramolecular, interpenetrando e fortalecendo os carboidratos estruturais, constituindo uma característica indesejável, responsável da recalcitrância da biomassa lignocelulósica para o seu uso na produção de E2G. Neste estudo, para compreender o processo da lignificação nas espécies de *Saccharum*, se

procedeu a determinar o conteúdo e composição (razão S/G) da lignina e analisar o perfil de ligações intermonoméricas. O conteúdo e composição da lignina pode variar em função de muitos fatores, tais como o tipo de tecido, idade da planta, genótipo, espécie e médio ambiente (de Souza et al., 2013).

A dosagem do conteúdo de lignina nas espécies de *Saccharum*, foi feito pelo método de Klason, que permite distinguir as frações solúvel e insolúvel em ácido sulfúrico, que somadas dão a quantidade de lignina total (Hatfield and Fukushima, 2005). Para todas as espécies em estudo, com relação à idade do internódio, se observa uma correlação negativa entre estes dois tipos de lignina Klason (Figuras 5A e 5B), indicando maior quantidade de lignina Klason solúvel nos internódios jovens e insolúvel, nos maduros. A explicação para isto é que a proção solúvel é formada principalmente por oligômeros em vias de polimerização, em um tecido que a formação da parede celular secundária ainda está ocorrendo em boa intensidade. O maior teor de lignina insolúvel nos maduros reflete a polimerização em estágio adiantado, e portanto, menos da fração solúvel é encontrada.

Os internódios maduros de *S. spontaneum* (20%) e *S. robustum* (18%) apresentaram maiores teores de lignina insolúvel quando comparados com aqueles de *S. barberi* (16%) e *S. officinarum* (14.5%). Os dados de lignina por Klason em internódios maduros de *S. spontaneum* e *S. robustum* foram corroborados pelos testes histoquímicos, nos quais a reação com fluoroglucinol foi notavelmente mais intensa nos elementos celulares destas espécies, apresentando paredes mais espessas e lignificadas (Figura 11).

A determinação da composição de lignina, expressada usualmente pela razão S/G, é um parâmetro pivotante, que prevê o grau e natureza da reticulação polimérica entre os monômeros que compõem a lignina, afetando por conseguinte o teor de recalcitrância na biomassa vegetal, desde que tecidos ricos em S são mais susceptíveis a degradação do que aqueles que são ricos em G (Ferrer et al., 2008). A proporção S/G afeta a degradabilidade da lignina e a produção de bioetanol lignocelulósico, desde que é um dos fatores principais intimamente relacionado com o rendimento da sacarificação. Uma significativa variação do coeficiente S/G, existe entre diferentes espécies de um mesmo gênero (Rencoret et al., 2008), cultivares (Santos et al., 2013), de planta em planta (Rodrigues et al., 1999), estádios de desenvolvimento (Galletti et al., 1996) e tecidos (Bottcher et al., 2013). Aqui neste estudo, a composição de lignina foi avaliada quantitativamente pelo método de digestão alcalina com NaOH e qualitativamente, por testes histoquímicos (Reação de Maule). É sabido que a proporção S/G aumenta progressivamente durante o desenvolvimento do caule e isto tem sido

apresentado para várias gramíneas tais como *Festuca arundinacea* (Chen et al., 2002), *Z. mays* (Jung and Casler, 2006), cana de açúcar (Bottcher et al., 2013) e *P. virgatum* (Shen et al., 2009). Nossos resultados mostram que nas espécies *S. officinarum* e *S. spontaneum* há tendência de aumento nos internódios maduros, mas não estatisticamente significativo. No entanto, muito diferentemente ao encontrado nas outras espécies, *S. barberi* e *S. robustum* apresentaram uma redução na relação S/G nos internódios maduros. estas observações foram confirmadas na análise histológica usando o teste histoquímico de Maüle (Figura 10), onde é possível observar que os elementos celulares dos internódios maduros correspondentes ao feixe vascular cortical e fibras do esclerênquima apresentam uma maior deposição de lignina G em *S. spontaneum* e *S. robustum* do que *S. officinarum* e *S. barberi*, e particularmente, nos feixes vasculares da medula de *S. spontaneum* e *S. robustum* começa a visualizar-se deposição de lignina G.

A conversão dos polímeros da parede celular dentro de açúcares fermentáveis, é um processo denominado sacarificação. Recalcitrância à sacarificação é a principal limitação para a conversão da biomassa lignocelulósica em etanol (Chen and Dixon, 2007). Dentro dos principais fatores que estão correlacionados com a recalcitrância à digestibilidade da celulose está a presença da lignina, representada pelo seu conteúdo-composição (Buxton and Russell, 1988), a arquitetura da parede celular e a estrutura anatômica dos tecidos influenciando significativamente na acessibilidade das enzimas hidrolíticas aos seus substratos (Wilson and Hatfield, 1997). Por outro lado, a susceptibilidade dos materiais lignocelulósicos à sacarificação vai ser de maneira diferente com relação ao estágio de maturidade e entre as espécies (Wilson and Hatfield, 1997). Neste estudo se observou que não houve diferença significativa na sacarificação nos internódios jovens quando comparadas a quatro espécies estudadas (Figura 6A). Este dado não teve correlação com a razão S/G, uma vez que maior S/G estaria relacionado à uma maior facilidade de sacarificação. Porém, uma vez que o teor de lignina solúvel indica ainda a deposição da lignina na parede celular secundária, a recalcitrância da parede ainda não seria significativa a ponto de inibir a sacarificação. De certa forma isto é embasado pelo fato de que internódios apresentam similares padrões de deposição de celulose, lignina e tipo de lignina, verificados histoquimicamente, pela reação com calcofluor (Figura 12), fluoroglucinol (Figura 11) e Maüle (Figura 10), respectivamente.

Por outro lado, maiores porcentagens de sacarificação foram observadas nos internódios maduros de *S. barberi* e *S. officinarum* do que *S. spontaneum* e *S. robustum*. A menor porcentagem de sacarificação em *S. robustum* e *S. spontaneum* é possível que seja uma

consequência da razão S/G nestas espécies, já que a primeira foi estatisticamente menor que as outras três espécies e a segunda apresentou tendências de ser menor que as outras duas espécies.

Limitações anatômicas que restringem a digestibilidade tem sido consideradas importantes desvantagens em caules maduros de gramíneas (Wilson and Hatfield, 1997). Células altamente lignificadas comprometem só uma pequena porção de seções transversais de caules, porém, devido a que as paredes dessas células são mais espessas que as do parênquima, a massa seca destas constituem uma significativa parte do caule (Chen et al., 2002). Fibras do esclerênquima devem a sua baixa digestibilidade tanto pelo seu alto conteúdo de lignina como por sua configuração anatômica (Grabber et al., 1992). A baixa porcentagem de sacarificação encontrada nos internódios maduros de *S. spontaneum* e *S. robustum* é possível que tenha uma explicação a partir de que em seções transversais dos internódios maduros destas espécies se observa que no córtex, elementos celulares tais como feixes vasculares, apresentam significativamente várias camadas de fibras do esclerênquima, espessas e mais lignificadas e com notável deposição de unidades G, quando comparando com os mesmos elementos celulares internódios maduros de *S. officinarum* e *S. barberi*.

A estrutura da lignina podem ser amplamente variáveis entre diferentes espécies de plantas e também entre os distintos tecidos dentro do mesmo organismo (del Río et al., 2012). A complexidade estrutural da lignina parece estar diretamente relacionada à frequência relativa das diferentes ligações interunidade, que por sua vez dependem da abundância de cada monômero durante o processo de polimerização. O estudo das ligações intermonoméricas e sua modificação pode ser explorados para a produção de ligninas mais degradáveis (Ralph et al., 1997; Grabber et al., 2008; Kiyota et al., 2012), permitindo o processamento mais eficiente da biomassa vegetal em direção aos açúcares fermentáveis na rota para a produção de bioetanol. As ligações 8-O-4( $\beta$  aril éter) são as mais comuns e caracterizadas como as de mais fácil clivagem. Ligninas ricas em unidades G, possuem mais ligações recalcitrantes, como 8-5 (fenilcoumaranas), 5-5 (resinóis) e 5-O-4, enquanto ligninas S, são menos interligadas e menos recalcitrantes a hidrólise (Kishimoto et al., 2009; Kiyota et al., 2012).

De forma geral, neste estudo a partir de um análise do perfil de oligômeros por UPLC/MS, foram identificados para as espécies de *Saccharum*, 11 estruturas, entre aldeídos, monômeros, dímeros e trímeros (Tabela 2). A distribuição desta estruturas permitiu uma clara distinção entre os internódios das espécies de *Saccharum*, a partir de que ocorreu uma maior

frequência dos oligômeros de lignina nos internódios maduros do que os jovens (Figura 7). Por outro lado, com relação à espécie, a maior frequência e diversidade dos oligômeros (dímeros e trímeros) de lignina foi encontrado de maneira notável nos internódios maduros *S. spontaneum* e *S. robustum* do que aqueles de *S. officinarum* e *S. barberi*. Morreel et al. (2010), afirma que a presença de vários oligômeros de lignina em tecido que sofrem extenso processo de lignificação são provenientes da disponibilidade de monolignóis que se acoplam em condição oxidativa para lignificação da parede celular, justificando a correlação entre teor de lignina e a frequência de oligômeros. Isto está em concordância com o encontrado em nosso estudo por que o conteúdo de lignina insolúvel foi maior para os internódios maduros de *S. spontaneum* e *S. robustum*. A ligação 8-O-4 foi o tipo mais frequente encontrado neste estudo, resultado que está em concordância com Santos et al. (2013), onde indica que esta estrutura é dominante nas ligninas de gramíneas correspondendo um 60 % do total. Outros trabalhos como os apresentados por Bottcher et al. (2013), corroboram também nossos resultados. Ligações neste estudo, contendo unidades G, apresentaram maior frequência do que as que contem unidades S. Não foi possível encontrar unidades H, apesar de que a lignina de gramíneas se caracterizam por possuir mais destas unidades (>5 %) do que a lignina de dicotiledôneas (< 5%) (Barrière et al., 2007). A justificativa de não ter identificado unidades H nas espécies de *Saccharum*, pode ser explicado a partir de que estas unidades ocorrem essencialmente como grupos fenólicos terminais livres, inertes, onde a sua incorporação detém o crescimento do polímero de lignina, devido ao seu alto potencial oxidativo, o que faz dessas unidades, sejam não reativas no acoplamento fenólico, sendo insolúveis no solvente acetato de etila usado neste experimento para a extração (Barrière et al., 2007). Ligações contendo só unidades G neste estudo, tais como os dímeros  $m/z$  357 G(8-5)G e  $m/z$  375 G(8-O-4)G e o trímero  $m/z$  553 G(8-O-4)G(8-5)G, foram identificadas notavelmente nos internódios maduros de *S. spontaneum* e *S. robustum*. Isto possivelmente explica que os internódios maduros destas espécies apresentaram uma razão S/G menor do que os internódios maduros de *S. officinarum* e *S. barberi*. Em concordância com o encontrado por (Marrone, 2015), As estruturas correspondentes aos dímeros  $m/z$  387 S(8-5)G e S(8-8)G e o trímero  $m/z$  583, G(8-O4)S(8-5)G foram identificados em todos os internódios das espécies em estudo, o que levaria a pensar que estas estruturas são conservadas e estariam cumprindo um papel importante no crescimento e desenvolvimento das espécies em estudo. O sinapil álcool (S) ( $m/z$ = 209) foi encontrado com maior frequência nos internódios jovens das espécies *S. officinarum* e *S. barberi* do que aqueles de *S. spontaneum* e *S. robustum*. Os

internódios jovens de *S. officinarum* e *S. barberi* apresentaram o maior conteúdo de lignina solúvel Klason do que as outras espécies de *Saccharum*, fato que poderia estar correlacionado com a maior frequência da estrutura  $m/z$  209, desde que as ligninas solúveis Klason estão constituídas maioritariamente de S (Yasuda et al., 2001). Entretanto comparando, nosso estudo com o perfil dos oligômeros de Bottcher et al. (2013), Marrone (2015) e Kiyota et al. (2012), apenas duas estruturas [ $m/z$  357, G(8-5)G;  $m/z$  583, G(8-O4)S(8-5)G] sempre estiveram presentes, indicando que são conservadas entre as espécies do gênero *Saccharum*.

Lignocelulose representa uma matéria prima renovável para a produção de biocombustíveis (Pauly and Keegstra, 2008). A O-acetilação dos polímeros da parede celular tem sido reconhecida a afetar negativamente o processamento da biomassa vegetal em etanol, desde que o acetato: (1) representa um potente inibidor para microrganismos fermentadores (Helle et al., 2003) e (2) gera um efeito estérico dificultando a hidrólise enzimática das hemiceluloses (Xiong et al., 2013). Portanto, uns dos focos de pesquisa que tem recebido atenção ultimamente, é a redução do conteúdo de grupos O-acetil em biomassa de espécies com características de bioenergia (Gille et al., 2011). As principais hemiceluloses em gramíneas são os xilanos (Pauly and Keegstra, 2008) e o seu grau de O-acetilação pode variar de acordo com as espécies, tipo de tecido e órgão e com o estado de desenvolvimento (Pauly and Keegstra, 2010). Acetilação de xilanos geralmente ocorre mais na posição C-2 do que C-3, mas a acetilação em ambas posições tem sido reportada (Bouveng, 1961; Lindberg et al., 1973). De acordo com a Figura 9D, se observa que a porcentagem total de acetilação (36.9% - 39.9%) nas espécies de *Saccharum* foi similar ao encontrado em outras biomassas de outras espécies (Pawar et al., 2013). Por outro lado, a acetilação na posição C-3 foi predominante (21.8%- 24.7%) com relação à substituição no C-2 (11.8%-13.0%) e a C-2/C-3 (1.47%- 3.47%). *S. robustum* e *S. spontaneum* foram as espécies que apresentaram os menores porcentagens de acetilação na posição C-3 (21.8% e 22.9%, respetivamente) e acetilação total (36.9%-37.4%, respetivamente). A hipótese de que, biomassas com reduzido porcentagens de acetilação, apresentam teores de sacarificação incrementados (Klein-Marcuschamer et al., 2010) aqui não foi evidenciada, desde *S. officinarum* e *S. barberi* com porcentagens de acetilação maiores que *S. spontaneum* e *S. robustum*, apresentaram maiores rendimentos de sacarificação. Desde que se conhece que nas paredes secundárias os xilanos se encontram estreitamente associados com a celulose (Scheller and Ulvskov, 2010b), uma menor porcentagem de grupos substituintes acetilos em *S. spontaneum* e *S. robustum* do que as outras espécies em estudo, poderia dar lugar a uma associação do xilano ainda mais estreita



com a celulose durante o processo de formação da parede celular, acrescentando a recalcitrância nestas espécies, limitando a enzimática sacarificação (Xiong et al., 2013).

A identificação e caracterização de sequências codificando as distintas enzimas catalíticas das etapas do metabolismo da lignina constitui um esforço inicial para se estudar e esclarecer esta rota metabólica em plantas dicotiledôneas e monocotiledôneas, permitindo o desenvolvimento de iniciativas na manipulação de teor de lignina e sua composição, a fim de melhorar a utilização da biomassa com características de bioenergia. As enzimas envolvidas nas reações da via biosintética da lignina são normalmente codificadas por famílias multigênicas para as distintas espécies de plantas (Bonawitz and Chapple, 2010). Porém, o número de isoformas presentes para cada gene influenciando no processo de lignificação pode variar entre uma particular espécie, órgão e estágio de desenvolvimento (Guillaumie et al., 2007).

Neste estudo a estratégia de se usar primers degenerados e sequenciar os amplicos era para exatamente identificar os genes mais expressos nas espécies estudadas, ainda que alguns genes fossem de família multigênicas. Ainda assim, para alguns dos genes foram isolados mais que um homólogo. Foram identificados 13 genes, dos quais CAD, CCoAOMT e 4CL, resultaram ser codificadas por mais de um gene.

No caso de 4CL apenas em *S. barberi* foram identificados dois genes, *Sh4CL1* e *Sh4CL3* (Bottcher et al., 2013), sendo que o primeiro representava 80% dos reads do sequenciamento, mostrando claramente codificar a isoformas predominante da enzima. Nas outras espécies foi identificado apenas *Sh4CL1*. Em híbridos de cana de açúcar foram identificados estes dois genes e também *Sh4CL2*, e nos tecidos novos e maduros o gene mais expresso foi *Sh4CL2*, seguido de *Sh4CL1* e *Sh4CL3*.

Para CCoAOMT foram achados dois genes para as quatro espécies. Porém, enquanto *ShCCoAOMT2* foi a mais expressa em *S. barberi*, ela foi menos expressa em *S. officinarum* e *S. robustum*, mas foi igualmente expressa a *ShCCoAOMT1* em *S. spontaneum*. Esses dois genes foram os dois mais expressos no cortex e medula de híbridos de cana de açúcar, sendo que houve variação de qual era o mais expresso dependendo da idade do internódio. Mas de modo geral *ShCCoAOMT1* foi mais expresso em tecidos maduros. Assim, de certa forma o comportamento contrastante entre as espécies aqui estudadas está *S. barberi*.

Para CAD foram identificados três genes, *ShCAD2*, *ShCAD7* e *ShCAD8*. Este último foi o gene mais expresso nas quatro espécies estudadas, sendo que em apenas *S. robustum* foram identificados os três genes. Nas outras três foram identificados *ShCAD2* e

*ShCAD8*. No estudo de Bottcher et al. (2013) claramente estes dois genes foram os mais expressos entre quatro identificados. *ShCAD7* foi mais expresso que *ShCAD6*, que não foi identificado aqui.

Para as outras enzimas da via de biossíntese de lignina foram identificados apenas um genes para cada enzima. Os genes aqui identificados foram aqueles mais expressos nos estudos de Bottcher et al. (2013). Esta informação em conjunto com o que foi comentado acima para 4CL, CCoAOMT e CAD mostram claramente que em *Sacharum* os genes mais expressos são os mesmos nas espécie (este estudo) e em híbridos delas (Bottcher et al., 2013), com pequena variação (CCoAOMT). Além disso, o alinhamento mostrou que existe grande conservação nas sequências, levando a concluir que apesar das duplicações no genoma da cana a rota se manteve conservada em termos dos principais genes contribuindo para a biossíntese de lignina.

## 7. CONCLUSÕES

Os resultados obtidos mostram que existem diferenças entre as espécies estudadas, seja em relação a carboidratos estruturais como também na quantidade e tipo de lignina. Os dados são coerentes, no sentido de que separam as duas canas que tem sido indicadas como cana energia - *S. spontaneum* e *S. robustum*, que acumulam mais fibra, se agruparam separadamente das outras duas, que acumulam mais açúcares solúveis na forma de sacarose. Além disso, as duas primeiras canas acumulam mais lignina insolúvel.

Entretanto, as informações de isolamento de genes mostrou uma grande conservação em termos de sequências, ou seja, a estratégia adotada mostrou que os genes principais na biossíntese de lignina nestas espécies são os mesmos e que as variações de conteúdo estariam relacionadas ao nível de expressão em cada espécie, informação que não foi obtida neste estudo. Assim, sendo as mesmas sequências, mas com supostos níveis de expressão diferentes, o controle da biossíntese de lignina em cana estaria a cargo de genes controladores, ou seja fatores de transcrição.

## MATERIAL SUPLEMENTAR FIGURAS SUPLEMENTARES

**Figuras S1-S10.** Análises filogenéticas de sequências de aminoácidos das enzimas da via biossintética dos monolignóis, obtidas da base de dados do NCBI, GeneBank, Phytozome e homólogos identificados neste estudo para as espécies de *Saccharum*. As anotações em azul mostram os genes (e respectivas SASs) identificados por Bottcher et al. (2013).

**Figura S1 - C4H**

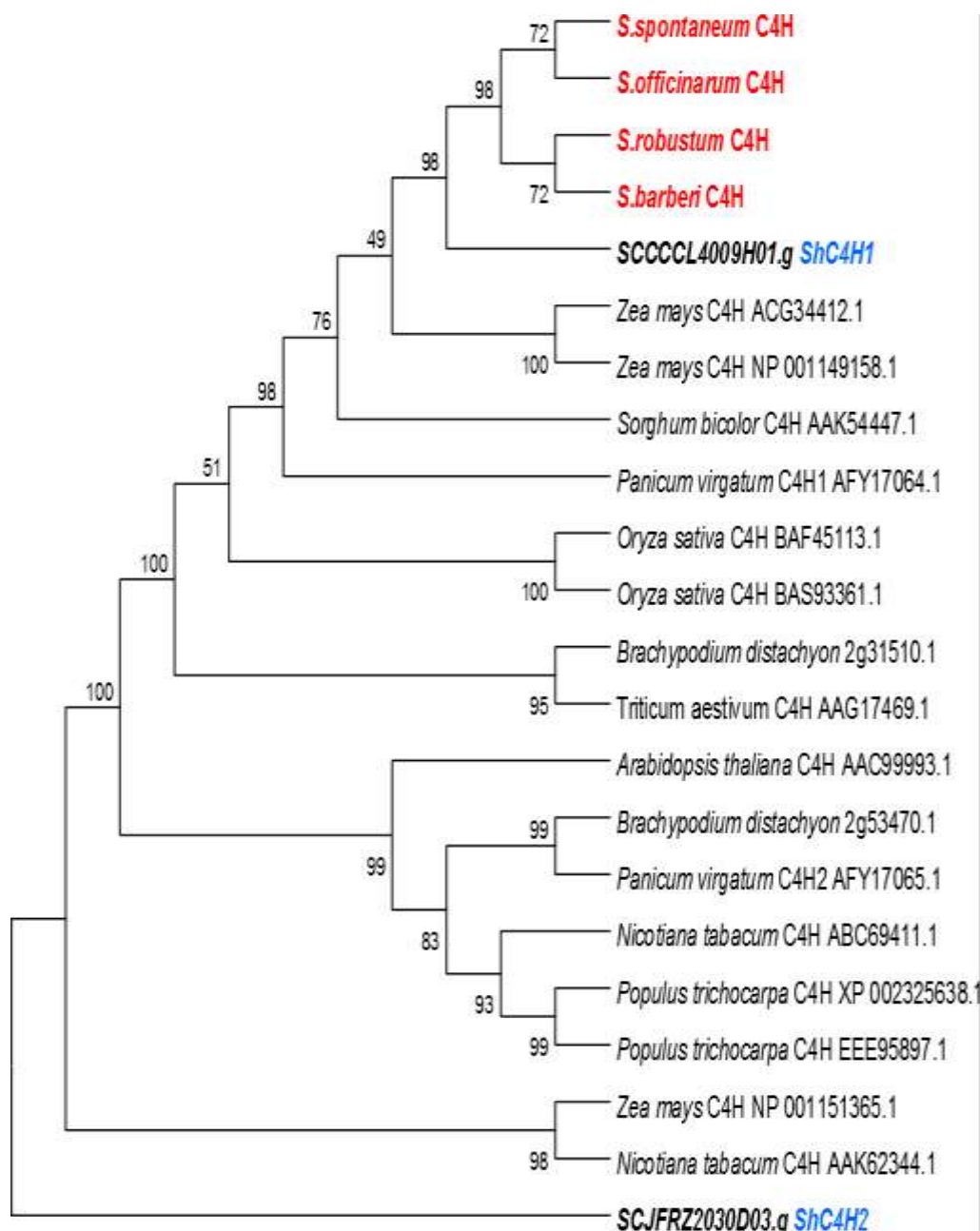
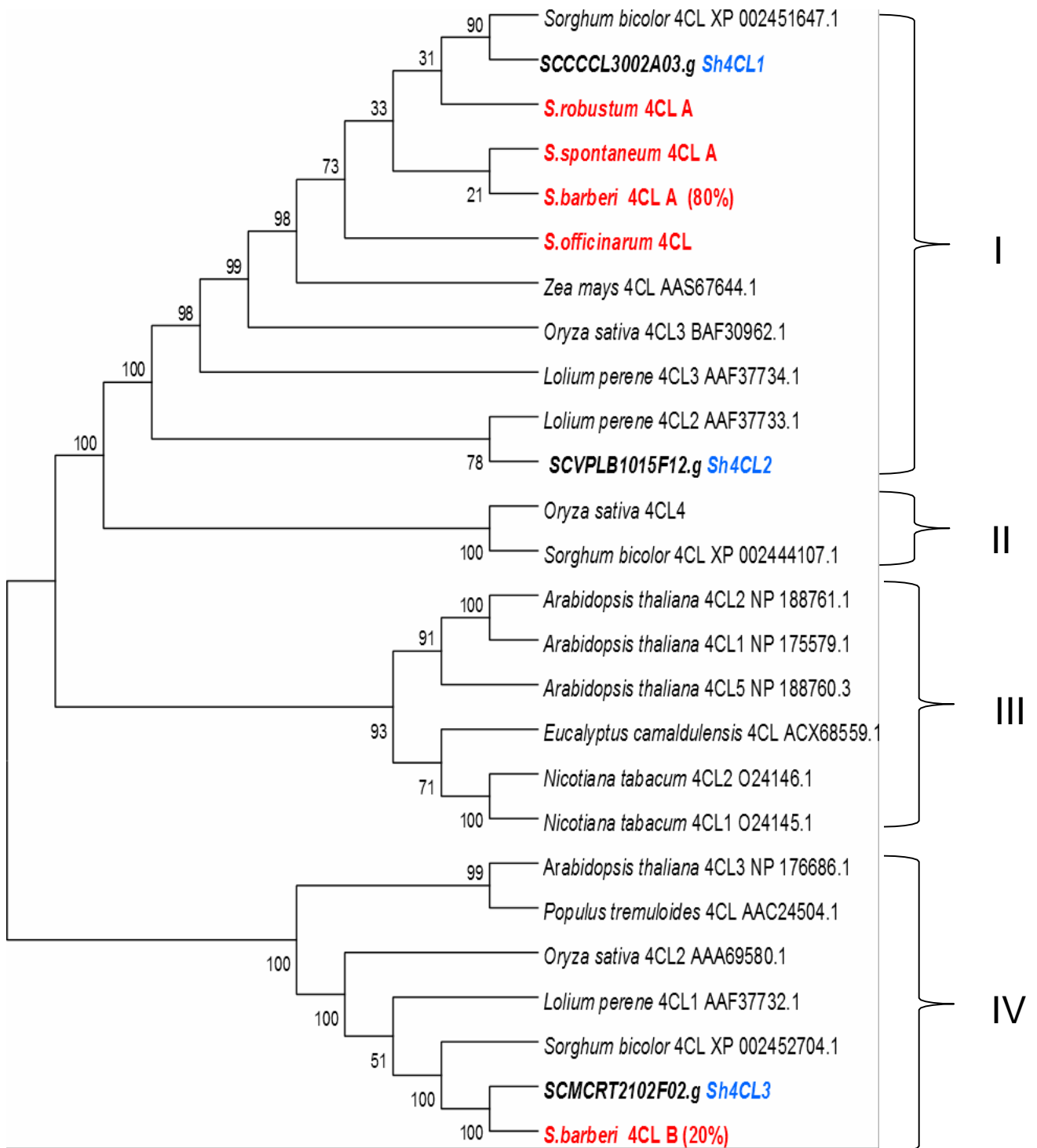


Figura S2 – 4CL



**Figura S3 – HCT**

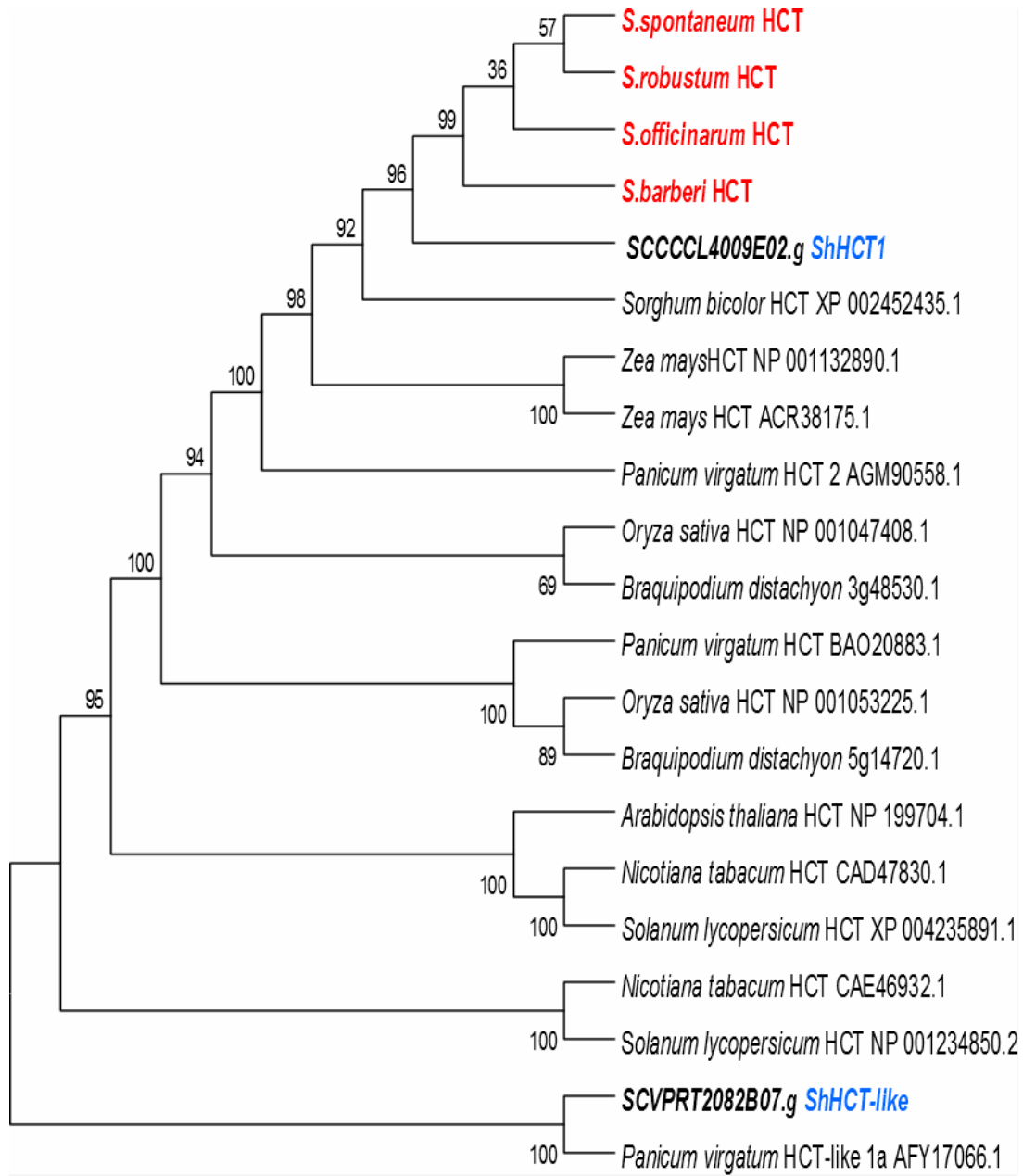


Figura S4 – C3H

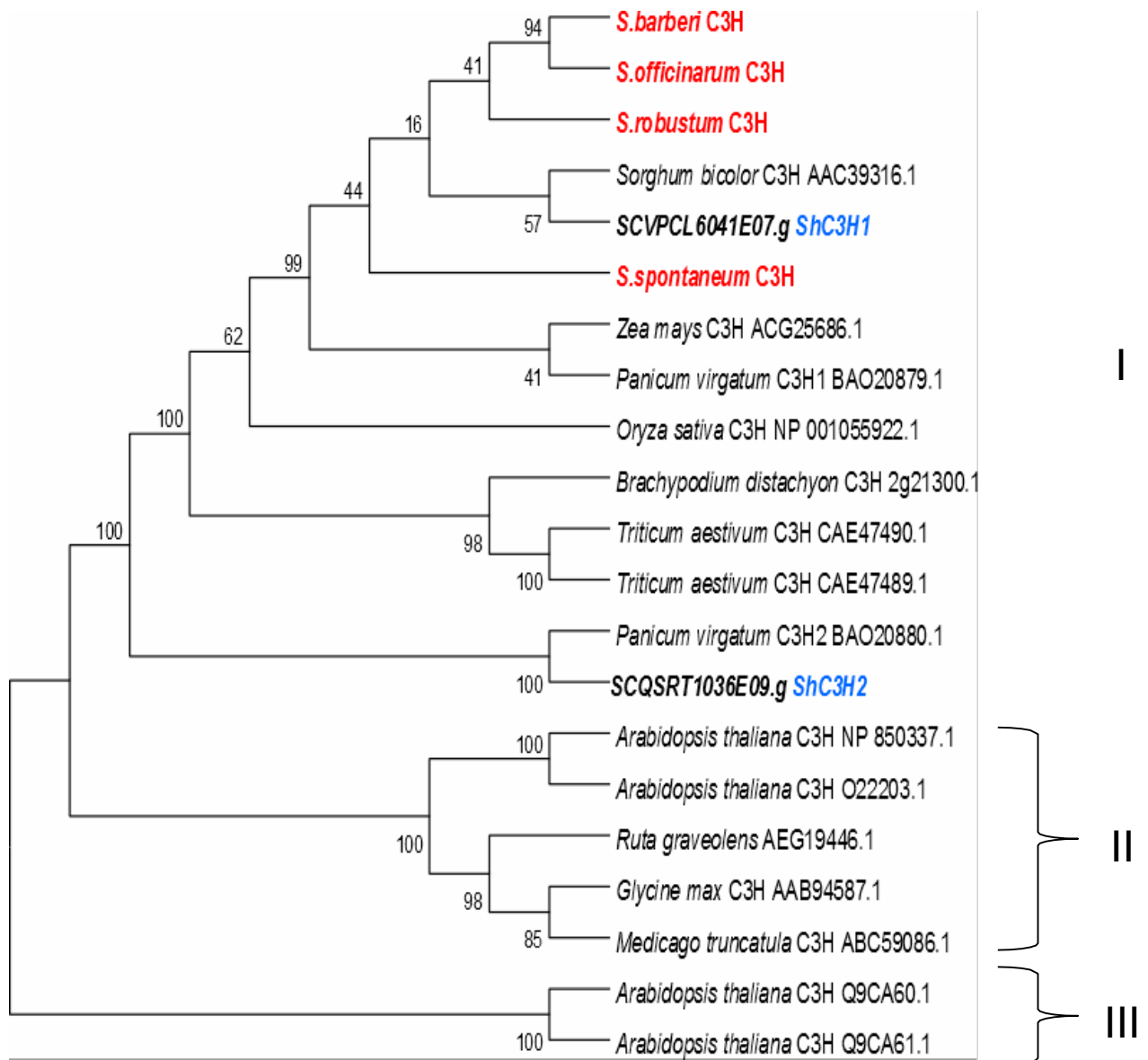


Figura S5 – CCoAOMT

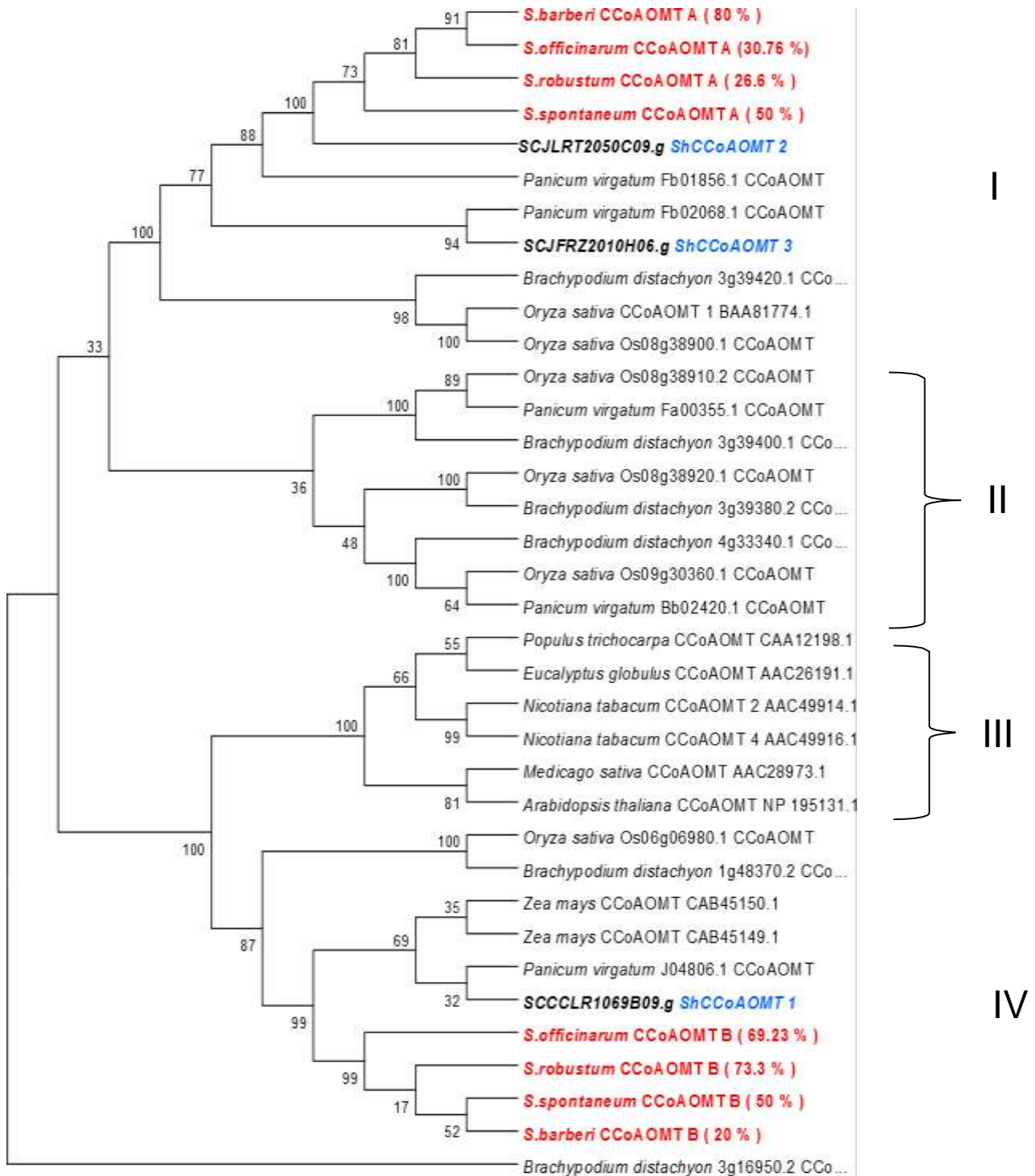




Figura S6 – CCR

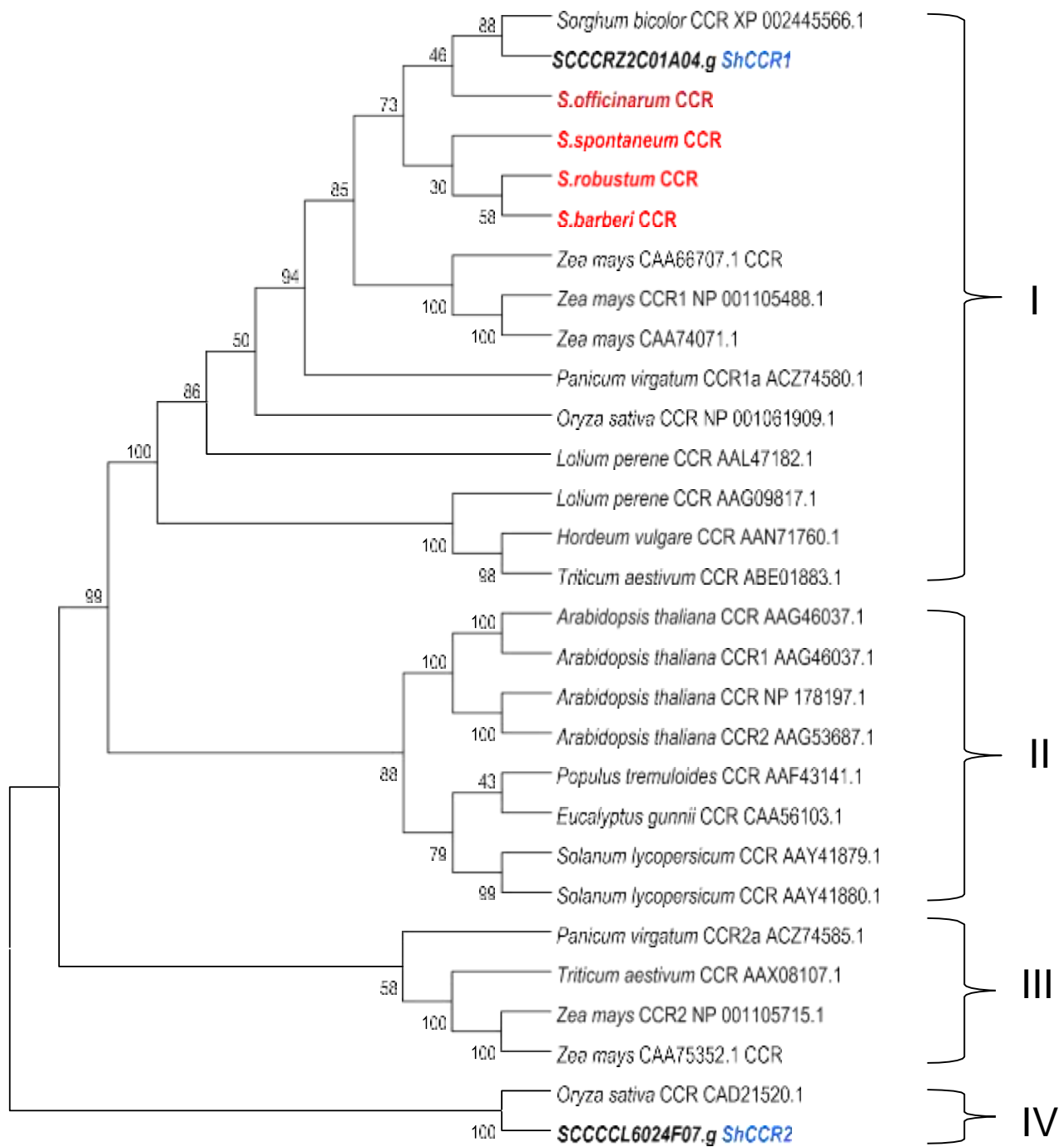


Figura S7 – F5H

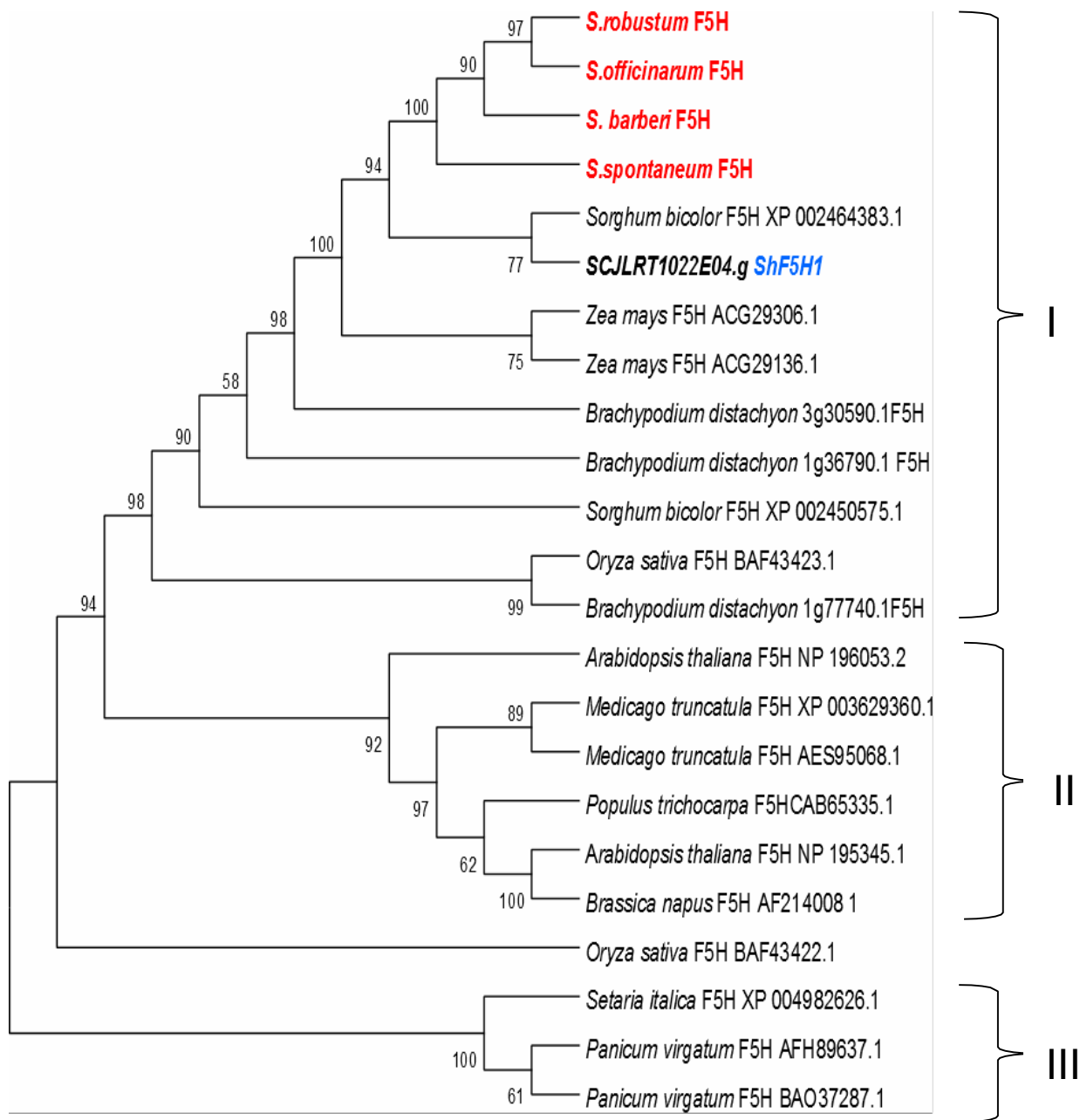


Figura S8 – COMT

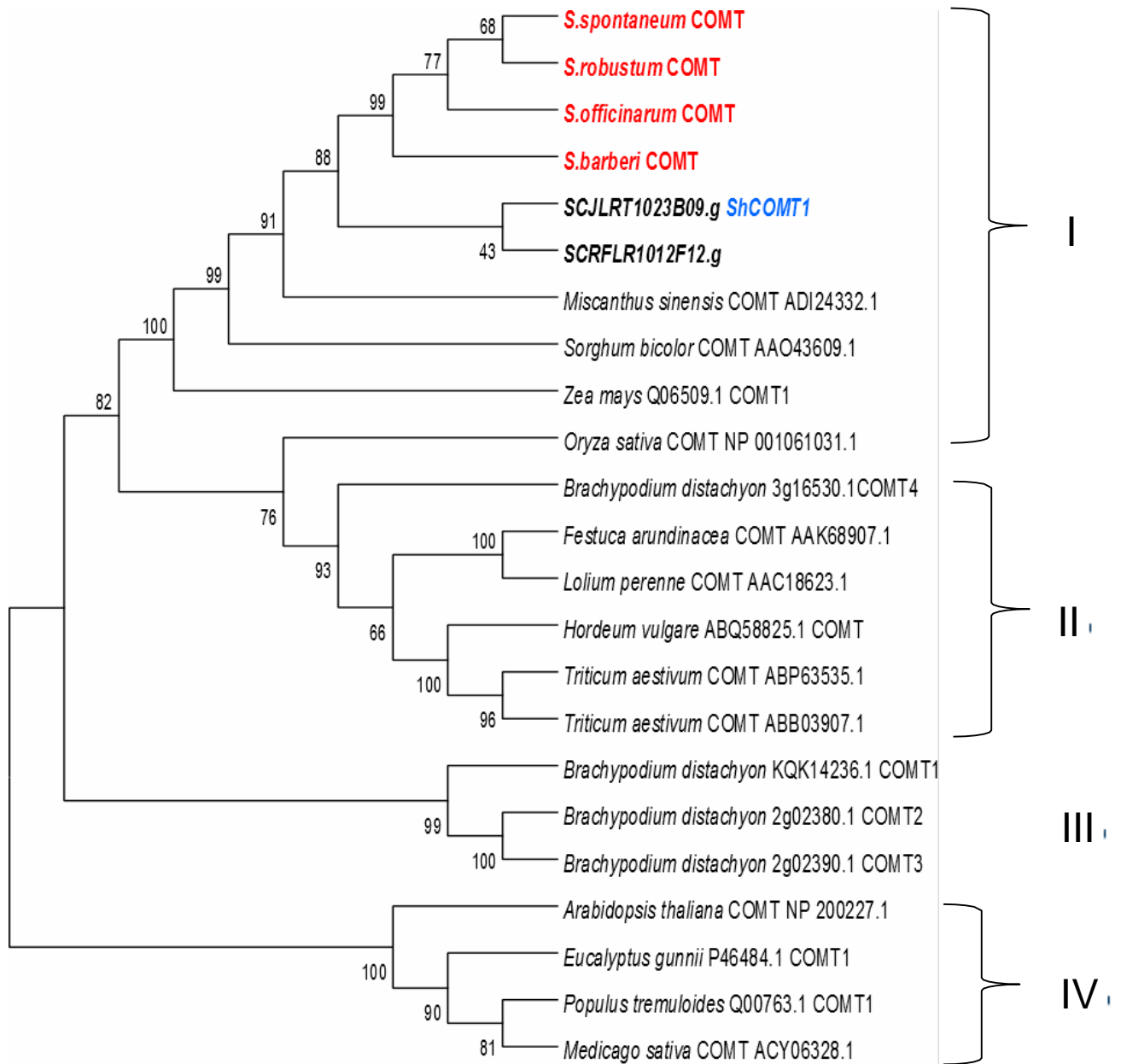
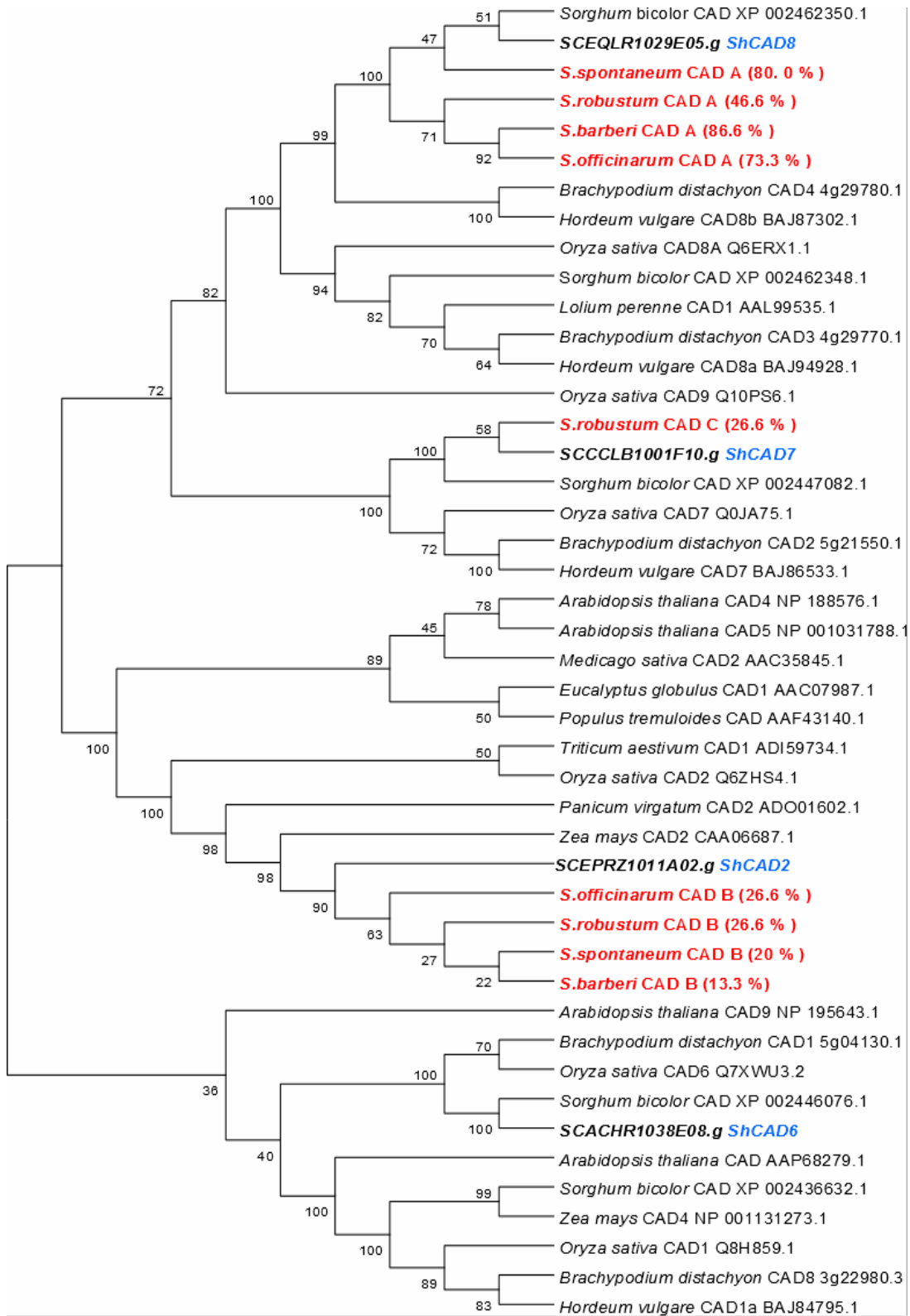


Figura S9 - CAD



**Figuras S1A-S10A. Alinhamento de seqüências de aminoácidos das enzimas da via bionssintética dos monolignóis nas espécies *Saccharum spontaneum* (Ss), *Saccharum barberi* (Sba), *Saccharum officinarum* (So), *Saccharum robustum* (Sr), *Miscanthus sinensis* (Ms), *Lolium perene* (Lp), *Panicum virgatum* (Pv) e *Sorghum bicolor* (Sb).**

**Figura S1A-C4H**

<i>ShC4H2</i> SCJFRZ2030D03.g	1	-----
Zm C4H	1	MDLALLEKALLGLFAAAVVAIAVAKLSGKRYRLPPGPPGAPVVGNNWLQVGD <sup>1</sup> DLNHRNLMA
Zm C4H	1	MDLALLEKALLGLFAAAVVAIAVAKLSGKRYRLPPGPPGAPVVGNNWLQVGD <sup>1</sup> DLNHRNLMA
<i>ShC4H1</i> SCCCCL4009H01.g	1	MDLVLLKALLGLFAAAVVAIAVAKLTGKRYRLPPGPPGAPVVGNNWLQVGD <sup>1</sup> DLNHRNLMA
*Ss C4H	1	-----LFGDIFL--RSTP
*So C4H	1	-----
*Sr C4H	1	-----
*Sba C4H	1	-----

<i>ShC4H2</i> SCJFRZ2030D03.g	1	-----
Zm C4H	61	MA <sup>1</sup> RFGDIFLLRMGVRNLVVVSTPELAKEVLHTQGVEFGSRTNRNVVDIFTGKGQDMVFT
Zm C4H	61	MA <sup>1</sup> RFGDIFLLRMGVRNLVVVSTPELAKEVLHTQGVEFGSRTNRNVVDIFTGKGQDMVFT
<i>ShC4H1</i> SCCCCL4009H01.g	61	LA <sup>1</sup> RFGDIFLLRMGVRNLVVVSTPELAKEVLHTQGVEFGSRTNRNVVDIFTGKGQDMVFT
*Ss C4H	12	RR <sup>1</sup> LFGDIFLLRMGVRNLVVVSTPELAKEVLHTQGVEFGSRTNRNVVDIFTGKGQDMVFT
*So C4H	1	---LFGDIFLLRMGVRNLVVVSTPELAKEVLHTQGVEFGSRTNRNVVDIFTGKGQDMVFT
*Sr C4H	1	---LFGDIFLLRMGVRNLVVVSTPELAKEVLHTQGVEFGSRTNRNVVDIFTGKGQDMVFT
*Sba C4H	1	---LFGDIFLLRMGVRNLVVVSTPELAKEVLHTQGVEFGSRTNRNVVDIFTGKGQDMVFT

<i>ShC4H2</i> SCJFRZ2030D03.g	1	-----
Zm C4H	121	VYGDHWRKMRRIMTVPF <sup>1</sup> TNKVVAQN <sup>1</sup> RAGWEEEARLVVEDVRKDPEAAAGGVVLRRLQL
Zm C4H	121	VYGDHWRKMRRIMTVPF <sup>1</sup> TNKVVAQN <sup>1</sup> RAGWEEEARLVVEDVRKDPEAAAGGVVLRRLQL
<i>ShC4H1</i> SCCCCL4009H01.g	121	VYGDHWRKMRRIMTVPF <sup>1</sup> TNKVVAQN <sup>1</sup> RAGWEEEARLVVEDVRDPRAAAEGVIRKRLQL
*Ss C4H	72	VYGDHWRKMRRIMTVPF <sup>1</sup> TNKVVAQN <sup>1</sup> RAGWEEEARLVVEDVRDPRAAAEGVIRKRLQL
*So C4H	58	VYGDHWRKMRRIMTVPF <sup>1</sup> TNKVVAQN <sup>1</sup> RAGWEEEARLVVEDVRDPRAAAEGVIRKRLQL
*Sr C4H	59	VYGDHWRKMRRIMTVPF <sup>1</sup> TNKVVAQN <sup>1</sup> RAGWEEEARLVVEDVRDPRAAAEGVIRKRLQL
*Sba C4H	58	VYGDHWRKMRRIMTVPF <sup>1</sup> TNKVVAQN <sup>1</sup> RAGWEEEARLVVEDVRDPRAAAEGVIRKRLQL

<i>ShC4H2</i> SCJFRZ2030D03.g	1	-MVKRMVC-AKAQ <sup>1</sup> VNKRWDEL <sup>1</sup> LKLIDDHANA <sup>1</sup> RS <sup>1</sup> ESQHVDEES <sup>1</sup> DFID <sup>1</sup> -----
Zm C4H	181	MMYNDMFRIMFDRRF <sup>1</sup> SE <sup>1</sup> HDPLFNK <sup>1</sup> LKALNAERSRL <sup>1</sup> SQSFEYNYGDFIPVLRPFLRGYLN
Zm C4H	181	MMYNDMFRIMFDRRF <sup>1</sup> SE <sup>1</sup> HDPLFNK <sup>1</sup> LKALNAERSRL <sup>1</sup> SQSFEYNYGDFIPVLRPFLRGYLN
<i>ShC4H1</i> SCCCCL4009H01.g	181	MMYNDMFRIMFDRRF <sup>1</sup> SE <sup>1</sup> EDPLFNK <sup>1</sup> LKALNAERSRL <sup>1</sup> SQSFEYNYGDFIPVLRPFLRGYLN
*Ss C4H	132	MMYNDMFRIMFDRRF <sup>1</sup> SE <sup>1</sup> EDPLFNK <sup>1</sup> LKALNAERSRL <sup>1</sup> SQSFEYNYGDFIPVLRPFLRGYLN
*So C4H	118	MMYNDMFRIMFDRRF <sup>1</sup> SE <sup>1</sup> EDPLFNK <sup>1</sup> LKALNAERSRL <sup>1</sup> SQSFEYNYGDFIPVLRPFLRGYLN
*Sr C4H	119	MMYNDMFRIMFDRRF <sup>1</sup> SE <sup>1</sup> EDPLFNK <sup>1</sup> LKALNAERSRL <sup>1</sup> SQSFEYNYGDFIPVLRPFLRGYLN
*Sba C4H	118	MMYNDMFRIMFDRRF <sup>1</sup> SE <sup>1</sup> EDPLFNK <sup>1</sup> LKALNAERSRL <sup>1</sup> SQSFEYNYGDFIPVLRPFLRGYLN

<i>ShC4H2</i> SCJFRZ2030D03.g	49	-----L <sup>1</sup> SVQOEYK <sup>1</sup> TRD <sup>1</sup> H <sup>1</sup> KAQ <sup>1</sup> VV <sup>1</sup> F
Zm C4H	241	RCHDLKTRRMK <sup>1</sup> EFEDNFVQERKKVMAQTGEIRCAMDHILEAERKGEINHDNVLYIVENIN
Zm C4H	241	RCHDLKTRRMK <sup>1</sup> EFEDNFVQERKKVMAQTGEIRCAMDHILEAERKGEINHDNVLYIVENIN
<i>ShC4H1</i> SCCCCL4009H01.g	241	RCHDLKTRRMK <sup>1</sup> VFEDNFVQERKKVMAQTGEIRCAMDHILEAERKGEINHDNVLYIVENIN
*Ss C4H	192	RCHNLKTRRMK <sup>1</sup> VFEDNFVQERKKVMAQTGEIRCAMDHILEAERKGEINHDNVLYIVENIN
*So C4H	178	RCHNLKTRRMK <sup>1</sup> VFEDNFVQERKKVMAQTGEIRCAMDHILEAERKGEINHDNVLYIVENIN
*Sr C4H	179	RCHDLKTRRMK <sup>1</sup> VFEDNFVQERKKVMAQTGEIRCAMDHILEAERKGEINHDNVLYIVENIN
*Sba C4H	178	RCHDLKTRRMK <sup>1</sup> VFEDNFVQERKKVMAQTGEIRCAMDHILEAERKGEINHDNVLYIVENIN

<i>ShC4H2</i> SCJFRZ2030D03.g	72	QA <sup>1</sup> TDSSFIVLE <sup>1</sup> TIKLMQNPNI <sup>1</sup> MIKLO <sup>1</sup> EF <sup>1</sup> RMTPK <sup>1</sup> KEFV <sup>1</sup> IEDDLNG <sup>1</sup> SYLK <sup>1</sup> AVIKE
Zm C4H	301	VAAIETTLWSIEWG <sup>1</sup> LAELVNHPAIQ <sup>1</sup> HKLREELASVLGAGV <sup>1</sup> PVTEPDLERLPYLQAI <sup>1</sup> VE
Zm C4H	301	VAAIETTLWSIEWG <sup>1</sup> LAELVNHPAIQ <sup>1</sup> HKLREELASVLGAGV <sup>1</sup> PVTEPDLERLPYLQAI <sup>1</sup> VE
<i>ShC4H1</i> SCCCCL4009H01.g	301	VAAIETTLWSIEWG <sup>1</sup> LAELVNHPAIQ <sup>1</sup> SKLREELASVLGAGV <sup>1</sup> PVTEPDLERLPYLQAI <sup>1</sup> VE
*Ss C4H	252	VAAIETTLWSIEWG <sup>1</sup> LAELVNHPAIQ <sup>1</sup> SKLREELASVLGAGV <sup>1</sup> PVTEPDLERLPYLQAI <sup>1</sup> VE
*So C4H	238	VAAIETTLWSIEWG <sup>1</sup> LAELVNHPAIQ <sup>1</sup> SKLREELASVLGAGV <sup>1</sup> PVTEPDLERLPYLQAI <sup>1</sup> VE
*Sr C4H	239	VAAIETTLWSIEWG <sup>1</sup> LAELVNHPAIQ <sup>1</sup> SKLREELASVLGAGV <sup>1</sup> PVTEPDLERLPYLQAI <sup>1</sup> VE
*Sba C4H	238	VAAIETTLWSIEWG <sup>1</sup> LAELVNHPAIQ <sup>1</sup> SKLREELASVLGAGV <sup>1</sup> PVTEPDLERLPYLQAI <sup>1</sup> VE

```

ShC4H2 SCJFRZ2030D03.g      132 TLRLHGPAPLLVPHLSMACDIEGYTIPSGTRVIVNAWALARDPTYWESASFMPERFLE
Zm C4H                      360 TLRLRMAIPLLVPHMNLNDGKLAGYDIPAESKILVNAWFLANDPKRWVRPDEFRPERFLE
Zm C4H                      360 TLRLRMAIPLLVPHMNLNDGKLAGYDIPAESKILVNAWFLANDPKRWVRPDEFRPERFLE
ShC4H1 SCCCCL4009H01.g      360 TLRLRMAIPLLVPHMNLNDGKLVGYDIPAESKILVNAWFLANDPKRWVRPDEFRPERFLE
*Ss C4H                     311 TLRLRMAIPLLVPHMNLNDGKLVGYDIPAESKILVNAWFLANDPKRWVRPDEFRPERFLE
*So C4H                     297 TLRLRMAIPLLVPHMNLNDGKLVGYDIPAESKILVNAWFLANDPKRWVRPDEFRPERFLE
*Sr C4H                     298 TLRLRMAIPLLVPHMNLNDGKLVGYDIPAESKILVNAWFLANDPKRWVRPDEFRPERFLE
*Sba C4H                    297 TLRLRMAIPLLVPHMNLNDGKLVGYDIPAESKILVNAWFLANDPKRWVRPDEFRPERFLE

ShC4H2 SCJFRZ2030D03.g      192 GGSAMTMDYRGNDFHYLPFGAGRRICPGTSFALSAETMLANLVYHFNWELPPELKKSGI
Zm C4H                      420 EEKS--VEAHGNDFRFVPFGVGRRSCPGIILALPIIGITLGRLVQNFQLLPPPGLDKIDT
Zm C4H                      420 EEKS--VEAHGNDFRFVPFGVGRRSCPGIILALPIIGITLGRLVQNFQLLPPPGLDKIDT
ShC4H1 SCCCCL4009H01.g      420 EEKT--VEAHGNDFRFVPFGVGRRSCPGIILALPIIGITLGRLVQNFQLLPPPGQDKIDT
*Ss C4H                     371 EEKT--VEAHGNDFRFVPFGVGRRSCPGIILALPIIGITLGRLVQNFHLLQSLV-----
*So C4H                     357 EEKT--VEAHGNDFRFVPFGVGRRSCPGIILALPIIGITLGRLVQNFHLLQSLV-----
*Sr C4H                     358 EEKT--VEAHGNDFRFVPFGVGRRSCPGIILALPIIGITLGRLVQNFHLLQSL-----
*Sba C4H                    357 EEKT--VEAHGNDFRFVPFGVGRRSCPGIILALPIIGITLGRLVQNFHLLQSLVQSL---

ShC4H2 SCJFRZ2030D03.g      252 DMTBSFGV---TMHRTEKLLVPLPQN
Zm C4H                      478 --TEKPSQFSNQLAKHATIVCKPEA--
Zm C4H                      478 --TEKPSQFSNQLAKHATIVCKPEA--
ShC4H1 SCCCCL4009H01.g      478 --TEKPSQFSNQLAKHATIVCKPEA--
*Ss C4H                      -----
*So C4H                      -----
*Sr C4H                      -----
*Sba C4H                      -----

```

# Figura S2A-4CL

```

Sh4CL3 SCMCR2102F02.g 1 MTIVAAPEAQPVAAAAVAPAAPEETVFRSKLPDIDIPSHLPHEHYCFARAAAYADAPCL
*Sba 4CL B 1 MTIVAAPEAQPVAAAAVAPAAPEETVFRSKLPDIDIPSHLPHEHYCFARAAAYADAPCL
Sb 4CL 1 MFSVDTAVAV-PVPVPEPEAEKEAAVFRSKLPDIDIPNSQSIDTYCEGKMSYADRACL
*Sr 4CL A 1 -----
*Sba 4CL A 1 -----
Sh4CL1 SCCCL3002A03.g 1 MGSVDTAVAV-PVPVAEPAAEEK-AVFRSKLPDIDIPNSQSIDHAYCEGKMSYADRACL
*Ss 4CL A 1 -----
*So 4CL A 1 -----
Sh4CL2 SCVPLB1015F12.g 1 -----
Lp 4CL 1 MGSTA-----ADAPPAEIVFRSKLPDIDIPTHLTIDQYCEQRLPISARACL

```

```

Sh4CL3 SCMCR2102F02.g 61 IATATGRITYTYETRLDRKAAASLHG GVGQGDVVMILLQNSVDFVLTFFGASFLGAVT
*Sba 4CL B 61 IATATGRITYTYETRLDRKAAASLHG GVGQGDVVMILLQNSVDFVLTFFGASFLGAVT
Sb 4CL 60 IDTGTCASYTYEVESSSRPAASGLRAGVGKGDVVMNLLRNCPFAFTFLGAARLGAAT
*Sr 4CL A 1 -----TSDFVFTFLGAARLGAAT
*Sba 4CL A 1 -----TSDFVFTFLGAARLGAAT
Sh4CL1 SCCCL3002A03.g 59 IDTGTCASYTYEVESSSRPAASGLRAGVGKGDVVMNLLRNCPFAFTFLGAARLGAAT
*Ss 4CL A 1 -----DLAFAFTFLGAARLGAAT
*So 4CL A 1 -----RFAFTFLGAARLGAAT
Sh4CL2 SCVPLB1015F12.g 1 -----
Lp 4CL 48 IDGATGAALTYEVDALSRRCAAGLRLGVCKGDVVMALLRNCPFEFAFVFLGAARLGAAT

```

```

Sh4CL3 SCMCR2102F02.g 121 TTANPFCTPLEIHKQFRASGSKLIVTQSAYVTKLRHEAFPRIGAASDGGEDEDNALTIT
*Sba 4CL B 121 TTANPFCTPLEIHKQFRASGSKLIVTQSAYVTKLRHEAFPRIGAASDGGEDEDNALTIT
Sb 4CL 120 TTANPFYTPHEIHRQAEAAAGAKVIVTEACAVEKVRFAAGR-----GVPVVT
*Sr 4CL A 19 TTANPFYTPHEIHRQAEAAAGAKLIVTEACAVEKVRFAAGR-----GVPVVT
*Sba 4CL A 19 TTANPFYTPHEIHRQAEAAAGAKLIVTEACAVEKVRFAAGR-----GVPVVT
Sh4CL1 SCCCL3002A03.g 119 TTANPFYTPHEIHRQAEAAAGAKLIVTEACAVEKVRFAAGR-----GVPVVT
*Ss 4CL A 20 TTANPFYTPHEIHRQAEAAAGAKLIVTEACAVEKVRFAAGR-----GVPVVT
*So 4CL A 17 TTANPFYTPHEIHRQAEAAAGAKLIVTEACAVEKVRFAAGR-----GTPVVT
Sh4CL2 SCVPLB1015F12.g 1 -----
Lp 4CL 108 TTANPFYTPHEIHRQATAAGARVIVTEACAVEKVRFAAER-----GTPVVS

```

```

Sh4CL3 SCMCR2102F02.g 181 IDDAANTPGCLAFVWLVTPPTDD-AALPVS SPDDFVALPSSGTTGLPKGVLTHTGGQ
*Sba 4CL B 181 IDDAANTPGCLAFVWLVTPADD-AALPVS SPDDFVALPSSGTTGLPKGVLTHTGGQ
Sb 4CL 167 VD--GRE-DGCVEFAEVIAAE---ELEADADVHPDDVVALPYSSGTTGLPKGVMLTHRSL
*Sr 4CL A 66 VD--GRE-DGCVEFAEVIAAE---ELEADADVHPDDVVALPYSSGTTGLPKGVMLTHRSL
*Sba 4CL A 66 VD--GRE-DGCVEFAEVIAAE---ELEADADVHPDDVVALPYSSGTTGLPKGVMLTHRSL
Sh4CL1 SCCCL3002A03.g 166 VD--GRE-DGCVEFAEVIAAE---ELEADADVHPDDVVALPYSSGTTGLPKGVMLTHRSL
*Ss 4CL A 67 VD--GRE-DGCVEFAEVIAAE---ELEADADVHPDDVVALPYSSGTTGLPKGVMLTHRSL
*So 4CL A 64 VD--GRE-DGCVEFAEVIAAE---ELEADADVHPDDVVALPYSSGTTGLPKGVMLTHRSL
Sh4CL2 SCVPLB1015F12.g 1 -----MDAAEPLADDEVDVDDVVALPYSSGTTGMPKGVMLTHRSL
Lp 4CL 155 VD--EGVDGCLIPFAETLLLEESGREFVDEAVDDVVALPYSSGTTGLPKGVMLTHRSL

```

```

Sh4CL3 SCMCR2102F02.g 240 VSNVAQQVDGENPNLYMREGDVALCLPLFHIISLNSVLLCLRAGAAVMLMKFEMGTM
*Sba 4CL B 240 VSNVAQQVDGENPNLYMREGDVALCLPLFHIISLNSVLLCLRAGAAVMLMKFEMGTM
Sb 4CL 221 ITSVAQQVDGENPNLYFSKDDVLLCLPLFHIISLNSVLLLAGLRAGSTIVIMRKFDLGAL
*Sr 4CL A 120 ITSVAQQVDGENPNLYFSKDDVLLCLPLFHIISLNSVLLLAGLRAGSTIVIMRKFDLGAL
*Sba 4CL A 120 ITSVAQQVDGENPNLYFSKDDVLLCLPLFHIISLNSVLLLAGLRAGSTIVIMRKFDLGAL
Sh4CL1 SCCCL3002A03.g 220 ITSVAQQVDGENPNLYFSKDDVLLCLPLFHIISLNSVLLLAGLRAGSTIVIMRKFDLGAL
*Ss 4CL A 121 ITSVAQQVDGENPNLYFSKDDVLLCLPLFHIISLNSVLLLAGLRAGSTIVIMRKFDLGAL
*So 4CL A 118 ITSVAQQVDGENPNLYFSKDDVLLCLPLFHIISLNSVLLLAGLRAGSTIVIMRKFDLGAL
Sh4CL2 SCVPLB1015F12.g 42 ITSVAQQVDGENPNLYFSKDDVLLCLPLFHIISLNSVLLLAGLRAGCAIVIMRKEIGAL
Lp 4CL 213 ITSVAQQVDGENPNLYFSKDDVLLCLPLFHIISLNSVLLLAGLRAGCAIVIMRKFDLGAL

```

```

Sh4CL3 SCMCR2102F02.g 300 LEHGTGDIG-----
*Sba 4CL B 300 LEGLQWRVTVAIVVPPAGCSRW--PR-----
Sb 4CL 281 VDLVRKHAITIAPFVPPIVVEIAKSPRVTAADLASIRVMMSGAAPMGKELQDAFMTKIPN
*Sr 4CL A 180 VDLVRKHAITIAPFVPPIVVEIAKSPRVTAADLASIRVMMSGAAPMGKELQDAFMTKIPN
Sba 4CL A 180 VDLVRKHAITIAPFVPPIVVEIAKSPRVTAADLASIRVMMSGAAPMGKELQDAFMTKIPN
Sh4CL1 SCCCL3002A03.g 280 VDLVRKHAITIAPFVPPIVVEIAKSPRVTAADLASIRVMMSGAAPMGKELQDAFMTKIPN
*Ss 4CL A 181 VDLVRKHAITIAPFVPPIVVEIAKSPRVTAADLASIRVMMSGAAPMGKELQDAFMTKIPN
*So 4CL A 178 VDLVRKHAITIAPFVPPIVVEIAKSPRVTAADLASIRVMMSGAAPMGKELQDAFMTKIPN
Sh4CL2 SCVPLB1015F12.g 102 VDLVRKHAITIAPFVPPIVVEIAKSPRVTAADLASIRVMMSGAAPMGKELQDAFMTKIPN
Lp 4CL 273 VDLVRKHITVAPFVPPIVVEIAKSPRVTAADLASIRVMMSGAAPMGKELQDAFMTKIPN

```

```

Sh4CL3 SCMCRT2102F02.g -----
*Sba 4CL B -----
Sb 4CL 341 AVLGGQGYGMTEAGPVLAMCLAFAKEPFQVKSGSGCTVVRNAELKVVD PDTGAALGRNQPG
*Sr 4CL A 240 AVLGGQGYGMTEAGPVLAMCLAFAKEPFQVKSGSGCTVVRNAELKIVDPDTGAALGRNQPG
*Sba 4CL A 240 AVLGGQGYGMTEAGPVLAMCLAFAKEPFQVKSGSGCTVVRNAELKIVDPDTGAALGRNQPG
Sh4CL1 SCCCL3002A03.g 340 AVLGGQGYGMTEAGPVLAMCLAFAKEPFQVKSGSGCTVVRNAELKIVDPDTGAALGRNQPG
*Ss 4CL A 241 AVLGGQGYGMTEAGPVLAMCLAFAKEPFQVKSGSGCTVVRNAELKIVDPDTGAALGRNQPG
*So 4CL A 238 AVLGGQGYGMTEAGPVLAMCLAFAKEPFQVKSGSGCTVVRNAELKIVDPDTGAALGRNQPG
Sh4CL2 SCVPLB1015F12.g 162 AVLGGQGYGMTEAGPVLAMCLAFAKEPFQVKSGSGCTVVRNAELKIVDPDTGASLGRNQPG
Lp 4CL 333 AVLGGQGYGMTEAGPVLAMCLAFAKEPFQVKSGSGCTVVRNAELKIVDPDTGASLGRNLPG

```

```

Sh4CL3 SCMCRT2102F02.g -----
*Sba 4CL B -----
Sb 4CL 401 EICIRGEQIMKGYLNDLESTKNTIDKDWLHTGDIGYVDDDDDEIFIVDRLKEIIKYKGFQ
*Sr 4CL A 300 EICIRGEQIMKGYLNDPESTKNTIDKDWLHTGDIGYVDDDDDEIFIVDRLKEIIKYKGFQ
*Sba 4CL A 300 EICIRGEQIMKGYLNDPESTKNTIDKDWLHTGDIGYVDDDDDEIFIVDRLKEIIKYKGFQ
Sh4CL1 SCCCL3002A03.g 400 EICIRGEQIMKGYLNDPESTKNTIDKDWLHTGDIGYVDDDDDEIFIVDRLKEIIKYKGFQ
*Ss 4CL A 301 EICIRGEQIMKGYLNDPESTKNTIDKDWLHTGDIGYVDDDDDEIFIVDRLKEIIKYKGFQ
*So 4CL A 298 EICIRGEQIMKGYLNDPESTKNTIDKDWLHTGDIGYVDDDDDEIFIVDRLKEIIKYKGFQ
Sh4CL2 SCVPLB1015F12.g 222 EICIRGEQIMKGYLNDPEATKNTIDKDWLHTGDIGYVDDDDDEIFIVDRLKEIIKYKGFQ
Lp 4CL 393 EICIRGEQIMKGYLNDPEATKNTIDKDWLHTGDIGYVDDDDDEIFIVDRLKEIIKYKGFQ

```

```

Sh4CL3 SCMCRT2102F02.g -----
*Sba 4CL B -----
Sb 4CL 461 VPPAELEALLITHPEIKDAAVVSMKDDLAGEIPVAFIVRTEGSEVTEDEIKQFVAKEVVF
*Sr 4CL A 360 VPPAELEALLITHPEIKDAAVVSMKDDLAGEIPVAFIVRTEGSEVTEDEIKQFVAKEVVF
*Sba 4CL A 360 VPPAELEALLITHPEIKDAAVVSMKDDLAGEIPVAFIVRTEGSEVTEDEIKQFVAKEVVF
Sh4CL1 SCCCL3002A03.g 460 VPPAELEALLITHPEIKDAAVVSMKDDLAGEIPVAFIVRTEGSEVTEDEIKQFVAKEVVF
*Ss 4CL A 361 VPPAELEALLITHPEIKDAAVVSMKDDLAGEIPVAFIVRTEGSEVTEDEIKQFVAKEVVF
*So 4CL A 358 VPPAELEALLITHPEIKDAAVVSMKDDLAGEIPVAFIVRTEGSEVTEDEIKQFVAKEVVF
Sh4CL2 SCVPLB1015F12.g 282 VPPAELEALLITHPEIKDAAVVSMKDELAGEVPVAFIIRSEGSEISENEIKQFVAKEVVF
Lp 4CL 453 VPPAELEALLITHPEIKDAAVVSMQDELAGEVPVAFVVRTEGSEISENEIKQFVAKEVVF

```

```

Sh4CL3 SCMCRT2102F02.g -----
*Sba 4CL B -----
Sb 4CL 521 YKRIHKVFFTESIPKNPSGKILRKDLRARLAAGVH-----
*Sr 4CL A 420 LQESL-----
*Sba 4CL A 420 YKNH-----
Sh4CL1 SCCCL3002A03.g 520 YKRVHKVFFTESIPKNPSGKILRKDLRARLAAGV-----
*Ss 4CL A 421 YKNHI-----
*So 4CL A 418 YKNH-----
Sh4CL2 SCVPLB1015F12.g 342 YKRINRVFFTDSIPKNPSGKILRKDLRARLAAGIPSSDNTQSKS
Lp 4CL 513 YKRICKVFFADSIPKSPSGKILRKDLRAKLAAGIPSSNTTQSKS

```



# Figura S3A-HCT

*Ss HCT	1	-----RIH--LGSEMVYPAAETPRRRLWNSGPDLVVP--RFHTPSVYFFRRRDADGNDLTA
Sb HCT	1	-----MKITVRGSEMVYPAAETPRRRLWNSGPDLVVP--RFHTPSVYFFRRRDADGNDLTA
ShHCT1 SCCCCL4009E02.g	1	-----MKITVRGSEMVYPAAETPRRRLWNSGPDLVVP--RFHTPSVYFFRRRDADGNDLTA
*So HCT	1	-----H--LGSEMVYPAAETPRRRLWNSGPDLVVP--RFHTPSVYFFRRRDADGNDLTA
*Sr HCT	1	-----H--LGSEMVYPAAETPRRRLWNSGPDLVVP--RFHTPSVYFFRRRDADGNDLTA
*Sba HCT	1	-----H--LGSEMVYPAAETPRRRLWNSGPDLVVP--RFHTPSVYFFRRRDADGNDLTA
ShHCT-like SCVPRT2082B07.g	1	MNTMDSEVQVVESSFVFNNEETPREGLMTSPIDLIVANFGIVYLVSYNNV-----A
Pv HCT-like 1a	1	-MKQAEVVVVVDLAIVPPSKETPGQPLWLSNLDLAVP--RHHTPLVYYYPAQAQG---AAA

*Ss HCT	53	PDGSFFDGARMRRALAEALVPFYPMAGRLARDE--DGRVEIDCNAAGVLFQEADAPDATID
Sb HCT	55	ADGSFFDGARMRRALAEALVPFYPMAGRLARDE--DGRVEIDCNAAGVLFQEADAPDATID
ShHCT1 SCCCCL4009E02.g	55	PDGSFFDGARMRRALAEALVPFYPMAGRLARDE--DGRVEIDCNAAGVLFQEADAPDATID
*So HCT	51	PDGSFFDGARMRRALAEALVPFYPMAGRLARDE--DGRVEIDCNAAGVLFQEADAPDATID
*Sr HCT	51	PDGSFFDGARMRRALAEALVPFYPMAGRLARDE--DGRVEIDCNAAGVLFQEADAPDATID
*Sba HCT	51	PDGSFFDGARMRRALAEALVPFYPMAGRLARDE--DGRVEIDCNAAGVLFQEADAPDATID
ShHCT-like SCVPRT2082B07.g	55	TADDFFDVARLKBAIAKALAAFYFPAAGRLGVNDDGRLEISCNTEGALFVVAHADDLVE
Pv HCT-like 1a	56	AGTGSFAPDRITAAALACALVPFYFPAAGRLSPGP--DGRPQINCTEGALFVVARADLTGDD

*Ss HCT	112	YFGDFAPTMEKRLIPTVDFSDDTAFPLLVLQVTHFKCGGVAIGVGMQHHVADGFGSGLHF
Sb HCT	114	YFGDFAPTMEKRLIPTVDFSDDTAFPLLVLQVTHFKCGGVAIGVGMQHHVADGFGSGLHF
ShHCT1 SCCCCL4009E02.g	114	YFGDFAPTMEKRLIPTVDFSDDTAFPLLVLQVTHFKCGGVAIGVGMQHHVADGFGSGLHF
*So HCT	110	YFGDFAPTMEKRLIPTVDFSDDTAFPLLVLQVTHFKCGGVAIGVGMQHHVADGFGSGLHF
*Sr HCT	110	YFGDFAPTMEKRLIPTVDFSDDTAFPLLVLQVTHFKCGGVAIGVGMQHHVADGFGSGLHF
*Sba HCT	110	YFGDFAPTMEKRLIPTVDFSDDTAFPLLVLQVTHFKCGGVAIGVGMQHHVADGFGSGLHF
ShHCT-like SCVPRT2082B07.g	115	DIKIFKFSFEIRLRIFVPIEFSSI---LIALQVTFKCGGVAIGTALFHVADLSSIFHF
Pv HCT-like 1a	115	IFEDFEPSPETLRFAFVPSRPFGDASCPLAVFQITFLKCGGVLTGTGTHHVLDDGVGDFQF

*Ss HCT	172	INSWADLCRGVPI---AVMPFIDRSLLRXXDPPTPAYPHIEYQPAPAMLSSEPPQAALT
Sb HCT	174	INSWADLCRGVPI---AVMPFIDRSLLRARDPPAFVYPHVEYQPAPAMLSSEPPQAALT
ShHCT1 SCCCCL4009E02.g	174	INSWADLCRGVPI---AVMPFIDRSLLRARDPPTPAYPHIEYQPAPAMLSSEPPQAALT
*So HCT	170	INSWADLCRGVPI---AVMPFIDRSLLRARDPPTPAYPHIEYQPAPAMLSSEPPQAALT
*Sr HCT	170	INSWADLCRGVPI---AVMPFIDRSLLRARDPPTPAYPHIEYQPAPAMLSSEPPQAALT
*Sba HCT	170	INSWADLCRGVPI---AVMPFIDRSLLRARDPPTPAYPHIEYQPAPAMLSSEPPQAALT
ShHCT-like SCVPRT2082B07.g	172	FKIWSAFSKHGDR--AAVEPCCHDRLLRARFPPTVHPDRT-----LDIAP-----
Pv HCT-like 1a	175	IQTWAGIARGIDAAEACSPFFHDXLLRAXVPAVEXLR-P-----XSSTPRXSAA

*Ss HCT	228	AKPATPAAAVAI FRLSRAELGRLRSQLPAR--EGAPRFSTYAVLAAHVRCASLARGLPA
Sb HCT	230	AKPATPPAAVAIFKL SRAELGRLRSQVPAREREGAPRFSTYAVLAAHVRCASLARGLPA
ShHCT1 SCCCCL4009E02.g	230	AKPATPPAAVAIFKLSRAELGRLRSQVPAR--EGAPRFSTYAVLAAHVRCASLARGLPA
*So HCT	226	AKPATPPAAVAIFKLSRAELGRLRSQVPAR--EDAPRFSTYAVLAAHVRCASLARGLPA
*Sr HCT	226	AKPATPPAAVAIFKLSRAELGRLRSQLPAR--EGAPGFSTYAVLAAHVRCASLARGLPA
*Sba HCT	226	AKPATPPAAVAIFKLSRAELGRLRSQVPAR--EGAPRFSTYAVLAAHVRCASLARGLPA
ShHCT-like SCVPRT2082B07.g		-----
Pv HCT-like 1a	226	---ARAEXVTRFIAISPKILADLNS-----RCAGVSTYCAVTAHLWRCVCVARGMAP

*Ss HCT	286	DQPTKLYCATDGRQRLQPPLPEGYFGNVIFTATPLANAGTVTA--GVAEGAGVIAAALDRM
Sb HCT	290	DQPTKLYCATDGRQRLQPPLPEGYFGNVIFTATPLANAGTVTA--GVAEGAAVIAAALDRM
ShHCT1 SCCCCL4009E02.g	288	DQPTKLYCATDGRQRLQPPLPEGYFGNVIFTATPLANAGTVTA--GVAEGAGVIAAALDRM
*So HCT	284	DQPTKLYCATDGRQRLQPPLPEGYFGNVIFTATPLANAGTVTA--GVAEGAGVIEALDRM
*Sr HCT	284	DQPTKLYCATDGRQRLQPPLPEGYFGNVIFTATPLANAGTVTA--GVAEGAGVIAAALDRM
*Sba HCT	284	DQPTKLYCATDGRQRLQPPLPEGYFGNVIFTATPLANAGTVTA--GVAEGAGVIAAALDRM
ShHCT-like SCVPRT2082B07.g		-----
Pv HCT-like 1a	276	GADTRLGLPANVRHRLSPPLPRSEFGNAVVRDLVTARVGDLLSPLGSAETIKKAVDGV

*Ss HCT	345	DDGYCRSALDYLELQPDLS-----ALVRGAHTFRCPNLGLTSWVRLPIHDADFVGWGRPVF
Sb HCT	349	DDGYCRSALDYLELQPDLS-----ALVRGAHTFRCPNLGLTSWVRLPIHDADFVGWGRPVF
ShHCT1 SCCCCL4009E02.g	347	DDGYCRSALDYLELQPDLS-----ALVRGAHTFRCPNLGLTSWVRLPIHDADFVGWGRPVF
*So HCT	343	DDGYCRSALDYLELQPDLS-----ALVRGAHTFRCPNLGLTSWVRLPIHDADFVGWGRPVF
*Sr HCT	343	DDGYCRSALDYLELQPDLS-----ALVRGAHTFRCPNLGLTSWVRLPIHDADFVGWGRPVF
*Sba HCT	343	DDGYCRSALDYLELQPDLS-----ALVRGAHTFRCPNLGLTSWVRLPIHDADFVGWGRPVF
ShHCT-like SCVPRT2082B07.g		-----
Pv HCT-like 1a	336	GDAFVRSVLDYLELEQAQKRGDDDSAQGEQMVPAITDLWAVSWIGMPPYSADFSGCAPRF

*Ss HCT	400	MGPGGIAYEGLAFVLPSANRDGSLSVAISLQAEHM-----
Sb HCT	404	MGPGGIAYEGLAFVLPSANRDGSLSVAISLQAEHMEKFRKLIYDF
ShHCT1 SCCCCL4009E02.g	402	MGPGGIAYEGLAFVLPSANRDGSLSVAISLQAEHMEKFRKLIYDF
*So HCT	398	MGPGGIAYEGLAFVLPSANRDGSLSVAISLQAEHMEKFRITS---
*Sr HCT	398	MGPGGIAYEGLAFVLPSANRDGSLSVAISLQAEHMEKFR-----
*Sba HCT	398	MGPGGIAYEGLAFVLPSANRDGSLSVAISLQAEHMEKFR-----
ShHCT-like SCVPRT2082B07.g		-----
Pv HCT-like 1a	396	VAPRQMGVGTAYMTFCANNDGRTVIFSMBAEHCECFEKVFGYGV

**Figura S4A- C3H**

*Ss C3H	1	-----KPIRCRCFQEWA
*Sr C3H	1	-----SINLRQIKPIRCRCFQEWA
*Sba C3H	1	-----INLRQIKPIRCRCFQEWA
*So C3H	1	-----INLRQIKPIRCRCFQEWA
ShC3H1 SCVPCL6041E07.g	1	MDASLLLSVALA-VVLIPLSLALLNRLRLGRLPPGPRPWPVLGNLRQIKPIRCRCFQEWA
Sb C3H	1	MDASLLLSVALA-VVLIPLSLALLNRLRLGRLPPGPRPWPVLGNLRQIKPIRCRCFQEWA
ShC3H2 SCQSRT1036E09.g	1	-----
Pv C3H	1	MNAASFLAVALALAALVPVSLLLNRLLYGKLPFGPRPRPVVGNLFDVQPVRCRCYQEWA
*Ss C3H	13	RYGPVISVWFGSGLTVVSTSELAKEVLKENDQQADRPNRNSTQRFNRNGQDLIWADY
*Sr C3H	20	RYGPVISVWFGSGLTVVSTSELAKEVLKENDQQADRPNRNSTQRFNRNGQDLIWADY
*Sba C3H	19	RYGPVISVWFGSGLTVVSTSELAKEVLKENDQQADRPNRNSTQRFNRNGQDLIWADY
*So C3H	19	RYGPVISVWFGSGLTVVSTSELAKEVLKENDQQADRPNRNSTQRFNRNGQDLIWADY
ShC3H1 SCVPCL6041E07.g	60	RYGPVISVWFGSGLTVVSTSELAKEVLKENDQQADRPNRNSTQRFNRNGQDLIWADY
Sb C3H	60	RYGPVISVWFGSGLTVVSTSELAKEVLKENDQQADRPNRNSTQRFNRNGQDLIWADY
ShC3H2 SCQSRT1036E09.g	1	-----MTVMLETSPTTVVSTSELAKEVLKTHDQQLADRCRDRSTESFSRGGQDLIWADY
Pv C3H	61	GRYGPIMTVWLGSGPTTVVSTAEELAREVLKTHDQSLADRVRRSRRERFSRGGKDLIWADY
*Ss C3H	73	
*Sr C3H	80	GPHYIKVRKLCNLELFTPKRLEALRPIREDEVTAMVESVYRAATAPGNEGKPLVVRNHLS
*Sba C3H	79	GPHYIKVRKLCNLELFTPKRLEALRPIREDEVTAMVESVYRAATAPGNEGKPLVVRNHLS
*So C3H	79	GPHYIKVRKLCNLELFTPKRLEALRPIREDEVTAMVESVYRAATAPGNEGKPLVVRNHLS
ShC3H1 SCVPCL6041E07.g	120	GPHYIKVRKLCNLELFTPKRLEALRPIREDEVTAMVESVYRAATAPGNEGKPLVVRNHLS
Sb C3H	120	GPHYIKVRKLCNLELFTPKRLEALRPIREDEVTAMVESVYRAATAPGNEGKPLVVRNHLS
ShC3H2 SCQSRT1036E09.g	55	GPHYIKVRKLCNLELFTQRLLEALRPIREDEVTAMVESVYRAATAPGNEGKPLVVRNHLS
Pv C3H	121	GAHYIKVRKLCNLELFTPRLLEALRPIREDEVTAMVESVYRAATAPGNEGKPLVVRNHLS
*Ss C3H	133	
*Sr C3H	140	MVAFNNITRLAFGKRFRMNANGDIDEQGREFKTIVNNGIKIGASLSVAEFIWYLRWLCPNL
*Sba C3H	139	MVAFNNITRLAFGKRFRMNANGDIDEQGEFETIVNNGIKIGASLSVAEFIWYLRWLCPNL
*So C3H	139	MVAFNNITRLAFGKRFRMNANGDIDEQGREFKTIVNNGIKIGASLSVAEFIWYLRWLCPNL
ShC3H1 SCVPCL6041E07.g	180	MVAFNNITRLAFGKRFRMNANGDIDEQGREFKTIVNNGIKIGASLSVAEFIWYLRWLCPNL
Sb C3H	180	MVAFNNITRLAFGKRFRMNANGDIDEQGREFKTIVNNGIKIGASLSVAEFIWYLRWLCPNL
ShC3H2 SCQSRT1036E09.g	115	MVAFNNITRLAFGKRFRVNAAGELG HNGIKIGASLSVAEFIWYLRWLCPNL
Pv C3H	181	MVAFNNITRLAFGKRFRVNAAGELG TDGIKIGASLSVAEFIWYLRWLCPNL
*Ss C3H	193	
*Sr C3H	200	EELYKTHNERRDRLTMKIIIEHAKSLKESGAKQHFDALFTLKEQYDLS EDTVIGLLWDM
*Sba C3H	199	EELYKTHNERRDRLTMKIIIEHAKSLKESGAKQHFDALFTLKEQYDLS EDTVIGLLWDM
*So C3H	199	EELYKTHNERRDRLTMKIIIEHAKSLKESGAKQHFDALFTLKEQYDLS EDTVIGLLWDM
ShC3H1 SCVPCL6041E07.g	240	EELYKTHNERRDRLTMKIIIEHAKSLKESGAKQHFDALFTLKEQYDLS EDTVIGLLWDM
Sb C3H	240	EELYKTHNERRDRLTMKIIIEHAKSLKESGAKQHFDALFTLKEQYDLS EDTVIGLLWDM
ShC3H2 SCQSRT1036E09.g	175	EQVFKAHGERRDRLTMKIIIEHAKSLKESGAKQHFDALFTLKEQYDLS EDTVIGLLWDM
Pv C3H	241	NEVFKAHGERRDRLTMKIIIEHAKSLKESGAKQHFDALFTLKEQYDLS EDTVIGLLWDM
*Ss C3H	253	
*Sr C3H	260	ITAGMDTTVISVEWAMAEVLVRNPRVQKKLQEEELDRVVGRDRVMLETDFQNLPLYLQAVVKE
*Sba C3H	259	ITAGMDTTVISVEWAMAEVLVRNPRVQKKLQEEELDRVVGRDRVMLETDFQNLPLYLQAVVKE
*So C3H	259	ITAGMDTTVISVEWAMAEVLVRNPRVQKKLQEEELDRVVGRDRVMLETDFQNLPLYLQAVVKE
ShC3H1 SCVPCL6041E07.g	300	ITAGMDTTVISVEWAMAEVLVRNPRVQKKLQEEELDRVVGRDRVMLETDFQNLPLYLQAVVKE
Sb C3H	300	ITAGMDTTVISVEWAMAEVLVRNPRVQKKLQEEELDRVVGRDRVMLETDFQNLPLYLQAVVKE
ShC3H2 SCQSRT1036E09.g	235	ITAGTDTTVISVEWAMAEVLVRNPRVQKKLQEEELDRVVGRDRVMLETDFQNLPLYLQAVVKE
Pv C3H	301	ITAGTDTTVISVEWAMAEVLVRNPRVQKKLQEEELDRVVGRDRVMLETDFQNLPLYLQAVVKE
*Ss C3H	313	
*Sr C3H	320	SLRLHPPTPLMLPHKASTNVKIGGYDIPKGANVMVNVWAVARDPKVWSNPLEYRPERFLE
*Sba C3H	319	SLRLHPPTPLMLPHKASTNVKIGGYDIPKGANVMVNVWAVARDPKVWSNPLEYRPERFLE
*So C3H	319	SLRLHPPTPLMLPHKASTNVKIGGYDIPKGANVMVNVWAVARDPKVWSNPLEYRPERFLE
ShC3H1 SCVPCL6041E07.g	360	SLRLHPPTPLMLPHKASTNVKIGGYDIPKGANVMVNVWAVARDPKVWSNPLEYRPERFLE
Sb C3H	360	SLRLHPPTPLMLPHKASTNVKIGGYDIPKGANVMVNVWAVARDPKVWSNPLEYRPERFLE
ShC3H2 SCQSRT1036E09.g	295	SLRLHPPTPLMLPHKASTNVKIGGYDIPKGANVMVNVWAVARDPKVWSNPLEYRPERFLE
Pv C3H	361	SLRLHPPTPLMLPHKASTNVKIGGYDIPKGANVMVNVWAVARDPKVWSNPLEYRPERFLE

*Ss C3H	373	ENIDIKGSDFRVL PFGAGRRVCPGAQLGINLVAS MIGHLLHHFEWSLPEGTRPEDVNMME
*Sr C3H	380	ENIDIKGSDFRVL PFGAGRRVCPGAQLGINLVAS MIGHLLHHFEWSLPEGTRPEDVNMME
*Sba C3H	379	ENIDIKGSDFRVL PFGAGRRVCPGAQLGINLVAS MIGHLLHHFEWSLPEGTRPEDVNMME
*So C3H	379	ENIDIKGSDFRVL PFGAGRRVCPGAQLGINLVAS MIGHLLHHFEWSLPEGTRPEDVNMME
ShC3H1 SCVPCL6041E07.g	420	ENIDIKGSDFRVL PFGAGRRVCPGAQLGINLVAS MIGHLLHHFEWSLPEGTRPEDINMME
Sb C3H	420	ENIDIKGSDFRVL PFGAGRRVCPGAQLGINLVAS MIGHLLHHFEWSLPEGTRPEDVNMME
ShC3H2 SCQSRT1036E09.g	355	ENIDIKGADFRVL PFGAGRRVCPGAQLGINLVAS MIGHMLHHFRWTLPEGTRPEDVSMME
Pv C3H	421	ENIDIKGADFRVL PFGAGRRVCPGAQLGINLV TMTLGHMLHHFSWSLPEGTRPEDVEMME
*Ss C3H	433	SPGLVTFMGTITQL-----
*Sr C3H	440	SPGLVTFMGTNH-----
*Sba C3H	439	SPGLVTFMGTITS-----
*So C3H	439	SPGLVTFMGTITS-----
ShC3H1 SCVPCL6041E07.g	480	SPGLVTFMGTPLQAVAKPRLEMEELYNRVPVEM
Sb C3H	480	SPGLVTFMGTPLQAVAKPRLEKEELYNRVPVEM
ShC3H2 SCQSRT1036E09.g	415	SPGLVTFMTPLQAVATPRLDKEELYRRVPSEI
Pv C3H	481	TPGLVTFMATPLKAVATPRLDREELYRRVPSEM

**Figura S5A-CCoAOMT**

```

Pv      CCoAOMT
ShCCoAOMT1 SCCCLR1069B09.g
*Sba    CCoAOMT B
*Ss     CCoAOMT B
*Sr     CCoAOMT B
*So     CCoAOMT B
Pv      CCoAOMT
*Ss     CCoAOMT A
ShCCoAOMT2 SCJLRT2050C09.g
*Sba    CCoAOMT A
*Sr     CCoAOMT A
*So     CCoAOMT A
ShCcoAOMT3 SCJFRZ2010H06.g
Pv      CCoAOMT

1 MASTAAEA--AKAAEQPANGNGEQKTRHSEVGHKSLKSDLYQYILDTSVYPREPESIK
1 MATTATEA--AKAA-PAEQANGEQKTRHSEVGHKSLKSDLYQYILDTSVYPREPESIK
1 ILKLCI-----QRVGSSPIWSTCRRPRIHFLKSEALYQYILDTSVYPREPESIK
1 -----TLYQYILDTSVYPREPESIK
1 -----LV-----QYILDTSVYPREPESIK
1 -----GNSILKSEALYQYILDTSVYPREPESIK
1 -----MAAGGDTIARVHTGIDSSNKLLKSEALYKYLDTSVLPHEPEICMR
1 -----LYKYVLDTSVLPHEPEICMR
1 -----MAPSGDTAIAQVHTGLDSSNKLLKSEALYKYLDTSVLPHEPEICMR
1 -----IQRVGS---SPIWSTCRRPRIHFLKSEDLKYKYLDTSVLPHEPEICMR
1 -----VLPHEPEICMR
1 -----AGGRGNYFLKSEDLKYKYLDTSVLPHEPEICMR
1 -----MDPGMATGGGGSIPDVHSNTDSSNKLLKSEQALYKYILDTSVLPNEPEICMR
1 MDTTASGRGGDPAMATAGGATAPNVHSNTDSSNKLLKSEQALYKYVLDTSVLPNEPEICMR

59 ELREITAKHPWNLMITSADEGQFLNMLIKLGAKTMEIGVYTGYSLLATALALPDGTH
58 ELREITAKHPWNLMITSADEGQFLNMLIKLGAKTMEIGVYTGYSLLATALALPDGTH
50 ELREITAKHPWNLMITSADEGQFLNMLIKLGAKTMEIGVYTGYSLLATALALPDGTH
21 ELREITAKHPWNLMITSADEGQFLNMLIKLGAKTMEIGVYTGYSLLATALALPDGTH
28 ELREITAKHPWNLMITSADEGQFLNMLIKLGAKTMEIGVYTGYSLLATALALPDGTH
29 ELREITAKHPWNLMITSADEGQFLNMLIKLGAKTMEIGVYTGYSLLATALALPDGTH
47 ELRLITDKHEWGLMQSSPDEAQLLRMLIKLGAARNTIEVGVTGYSLLATALALPDGKV
20 ELRLITDKHEWGFQSSPDEAQLLRMLIKLGAARNTIEVGVTGYSLLATALALPDGKV
48 ELRLITDKHEWGFQSSPDEAQLLRMLIKLGAARNTIEVGVTGYSLLATALALPDGKV
45 ELRLITDKHEWGFQSSPDEAQLLRMLIKLGAARNTIEVGVTGYSLLATALALPDGKV
11 ELRLITDKHEWGFQSSPDEAQLLRMLIKLGAARNTIEVGVTGYSLLATALALPDGKV
33 ELRLITDKHEWGFQSSPDEAQLLRMLIKLGAARNTIEVGVTGYSLLATALALPDGKV
52 ELRLITDKHERRYATPPDEAQLLRMLIKLGAARNTIEVGVTGYSLLATALALPDGKV
61 ELRLITDKHERRNMTAPDEAQLLRMLIKLGAARNTIEVGVTGCSLISTATALALPDGKV

119 LAMDINRENYELGFCIEKAGVAHKIDFREGPALFVLDDLIADENHGTFFDAFVADAKD
118 LAMDINRENYELGFCIEKAGVAHKIDFREGPALFVLDDLIADENHGSFDFVFVADAKD
110 LAMDINRENYELGFCIEKAGVAHKIDFREGPALFVLDDLIADENHGSFDFVFVADAKD
81 LAMDINRENYELGFCIEKAGVAHKIDFHEGPALFVLDDLIADENHGSFDFVFVADAKD
88 LAMDINRENYELGFCIEKAGVAHKIDFREGPALFVLDDLIADENHGSFDFVFVADAKD
89 LAMDINRENYELGFCIEKAGVAHKIDFREGPALFVLDDLIADENHGSFDFVFVADAKD
107 LALVDVREYYEIGRPFIEKAGVAGKVDREGPALDGLDALLADEANLGAFDFAFVADAKP
80 LAFDVSREYYDIEGRPFIEKAGVAHKVDREGPALEGLDALLADEANHGAFDFAFVADAKP
108 LAFDVSREYYDIEGRPFIDKAGVAHKVDREGPALEGLDALLADEANHGAFDFAFVADAKP
105 LAFDVSREYYDIEGRPFIEKAGVAHKVDREGPALEGLDALLADEANHGAFDFAFVADAKP
71 LAFDVSREYYDIEGRPFINKAGVAHKVDFRGGPALEGLDALLADEANHGAFDFAFVADAKP
93 LAFDVSREYYDIGRPFIEKAGVAHKVDREGPALEGLDALLADEANHGAFDFAFVADAKP
112 LAFDVSREYYELGPFIEKAGVAHKVDREGPALEGLDALLADEANHGAFDFAFVADAKP
121 LAFDVSREYYEIGRPFIEKAGVAHKVDREGPALERHEGLLADANLGAFDFAFVADAKP

179 NYLNYHERLLLVKLGGIGYDNTLWNGSVLPDDAPMRKYIRFRDFVLVNKAFAADE
178 NYLNYHERLLLVKLGGIGYDNTLWNGSVLPDDAPMRKYIRFRDFVLVNKAFAADE
170 NYLNYHERLLLVKLGGIGYDNTLWNGSVLPDDAPMRKYIRFRDFVLVNKAFAADE
141 NYLNYHERLLLVKLGGIGYDNTLWNGSVLPDDAPMRKYIRFRDFVLVNKAFAADE
148 NYLNYHERLLLVKLGGIGYDNTLWNGSVLPDDAPMRKYIRFRDFVLVNKAFAADE
149 NYLNYHERLLLVKLGGIGYDNTLWNGSVLPDDAPMRKYIRFRDFVLVNKAFAADE
167 NYVVGGSIVYDNTLWAGTVALPPDTPMSDRRFSAAVRDLNVRLSKDE
140 NYVVGGSIVYDNTLWAGTVALPPDTPMNDLDRRFSAAVRDLNVRLSKDE
168 NYVVGGSIVYDNTLWAGTVALPPDTPMNDLDRRFSAAVRDLNVRLSKDE
165 NYVVGGSIVYDNTLWAGTVALPPDTPMNDLDRRFSAAVRDLNVRLSKDE
131 NYVVGGSIVYDNTLWAGTVALPPDTPMNDLDRRFSAAVRDLNVRLSKDE
153 NYVVGGSIVYDNTLWAGTVALPPDTPMNDLDRRFSAAVRDLNVRLSKDE
172 NYVVGGAIVYDNTLWDGTVALPPDAPMSDRRFSAAVRDLNVRLSKDE
181 NYVVGGAIVYDNTLWDGTVALPPDAPMSDRRFSAAVRDLNVRLSKDE

```

Pv CCoAOMT  
*ShCCoAOMT1* SCCCLR1069B09.g  
 \*Sba CCoAOMT B  
 \*Ss CCoAOMT B  
 \*Sr CCoAOMT B  
 \*So CCoAOMT B  
 Pv CCoAOMT  
 \*Ss CCoAOMT A  
*ShCCoAOMT2* SCJLRT2050C09.g  
 \*Sba CCoAOMT A  
 \*Sr CCoAOMT A  
 \*So CCoAOMT A  
*ShCCoAOMT3* SCJFRZ2010H06.g  
 Pv CCoAOMT

239 RVEICQLPVGD~~SVT~~CR~~R~~K-----  
 238 RVEICQLPVGD~~SVT~~CR~~R~~KKTC--  
 230 RVEVCQLNH~~P~~AAAG~~EHM~~----  
 201 RVEICQLNH-----  
 208 RVEICQLNH~~R~~RLQVDH~~G~~----  
 209 RVEICQLNH~~R~~RLQVDH~~G~~EL---  
 227 RVEVCQLAIAD~~SVT~~CR~~R~~V-----  
 200 RVEICQLKS~~VN~~-----  
 228 RVEVCQLAIAD~~SVT~~CR~~R~~V-----  
 225 RVEICQLNH~~R~~RLQVDH~~G~~ELPTR  
 191 RVEVCQLNH-----  
 213 RVEICQLNH~~R~~RLQVDH~~G~~ELPT-  
 232 RVE~~I~~COLAVAD~~SVT~~CR~~R~~V-----  
 241 RVE~~I~~COLTVAD~~SVT~~CR~~R~~V-----

Figura S6A-CCR

```

Sb CCR 1 MTVVDAVSTDAAGAAPAA--AAAPVVVAQPGNGQTVVCVTGAGGYIASWLVKMLLEKGYTV
ShCCR1 SCCRZ2C01A04.g 1 MTIVDAVSTDAAGAPAA-----AAALVQPAGNGQTVVCVTGAGGYIASWLVKLLLEKGYTV
*Sr CCR 1 -----LVVICSWVVKLLLEKGYTV
*Sba CCR 1 -----H--FESWVVKLLLEKGYTV
*Ss CCR 1 -----FASWVVKLLLEKGYTV
*So CCR 1 -----IASWVVKLLLEKGYTV
ShCCR2 SCCCCL6024F07.g 1 MA-----PPPSVGVLRFAYWVSPL-HLGNSNQQLVVCVTGAGGFIQSWVVKELLQRGYRV
Os CCR 1 -----MSSNFEANNNGEKQLVVCVTGAGGFIQSWVVKELLIRGYHV

Sb CCR 59 KGTVRNPDDPKNAHLKALDGAERLILCKADLLDYDAICRAVQGCQG VFHTASPVTDDE
ShCCR1 SCCRZ2C01A04.g 56 KGTVRNPDDPKNAHLKALDGAERLILCKADLLDYDAICRAVQGCCHGVFTASPVTDDE
*Sr CCR 19 KGTVRNPDDPKNAHLKALDGAERLILCKADLLDYDAICRAVQGCCHGVFTASPVTDDE
*Sba CCR 18 KGTVRNPDDPKNAHLKALDGAERLILCKADLLDYDAICRAVQGCCHGVFTASPVTDDE
*Ss CCR 17 KGTVRNPDDPKNAHLKALDGAERLILCKADLLDYDAICRAVQGCCHGVFTASPVTDDE
*So CCR 17 KGTVRNPDDPKNAHLKALDGAERLILCKADLLDYDAICRAVQGCCHGVFTASPVTDDE
ShCCR2 SCCCCL6024F07.g 55 GTA KNAHL LALGAKERLT CRADVLDAS HA AGCHGVFHVASPVNDPE
Os CCR 42 GTARDPDDPKNAHLLELTGADQRLSLCRADVLDAS RA AFSGCHGVFHVASPVNDPE

Sb CCR 119
ShCCR1 SCCRZ2C01A04.g 116 QMVEPAVRGTEYVINAAAEAGTVRRVFTSSIGAVTMDPSRGPDVVVDESCWSDLEFCCK
*Sr CCR 79 QMVEPAVRGTEYVINAAAEAGTVRRVFTSSIGAVTMDPSRGPDVVVDESCWSDLEFCCK
*Sba CCR 78 QMVEPAVRGTEYVINAAAEAGTVRRVFTSSIGAVTMDPSRGPDVVVDESCWSDLEFCCK
*Ss CCR 77 QMVEPAVRGTEYVINAAAEAGTVRRVFTSSIGAVTMDPSRGPDVVVDESCWSDLEFCCK
*So CCR 77 QMVEPAVRGTEYVINAAAEAGTVRRVFTSSIGAVTMDPSRGPDVVVDESCWSDLEFCCK
ShCCR2 SCCCCL6024F07.g 115 L-VPVAVEGTRNVINVAALLEG-ARRVFTSSYGAVHMDPSRSPDAVLDEICWSDYEFCKR
Os CCR 102 L-VPVAVEGTRNVINAAALMG-VRRVFTSSYGAVHMDPSRSPDAVLDEICWSDYEFCKR

Sb CCR 179 TRNWCYCYGKAVAEQAAWDAARQGVDLVVVNPVLVVGPLLQPTVNASIAHVLYLDGSR
ShCCR1 SCCRZ2C01A04.g 176 TRNWCYCYGKAVAEQAAWDAARQGVDLVVVNPVLVVGPLLQPTVNASIAHVLYLDGSR
*Sr CCR 139 TRNWCYCYGKAVAEQAAWDAARQGVDLVVVNPVLVVGPLLQPTVNASIAHVLYLDGSR
*Sba CCR 138
*Ss CCR 137 TRNWCYCYGKAVAEQAAWDAARQGVDLVVVNPVLVVGPLLQPTVNASIAHVLYLDGSR
*So CCR 137 TRNWCYCYGKAVAEQAAWDAARQGVDLVVVNPVLVVGPLLQPTVNASIAHVLYLDGSR
ShCCR2 SCCCCL6024F07.g 173 TDNLYCCAKMMAETITATEAAAGLQIADVLPCTMGPLLQQTINISNYHVAAYVMGTR
Os CCR 160 TDNLYCCAKMMAEMTATEAAAGLQIADVLPCTMGPLLQQTINISNYHVAAYVMGTR

Sb CCR 239 TFANAVQAYVDVRDVAHAHLRVFESPAASGRYLCAERVLHREDVVRI LAKLFPEYPVPTR
ShCCR1 SCCRZ2C01A04.g 236 TFANAVQAYVDVRDVAHAHLRVFESPRASGRYLCAERVLHREDVVRI LAKLFPEYPVPTR
*Sr CCR 199 TFANAVQAYVDVRDVAHAHLRVFESPRASGRYLCAERVLHREDVVRI LAKLFPEYPVPTR
*Sba CCR 198 TFANAVQAYVDVRDVAHAHLRVFESPRASGRYLCAERVLHREDVVRI LAKLFPEYPVPTR
*Ss CCR 197 TFANAVQAYVDVRDVAHAHLRVFESPRASGRYLCAERVLHREDVVRI LAKLFPEYPVPTR
*So CCR 197 TFANAVQAYVDVRDVAHAHLRVFESPRASGRYLCAERVLHREDVVRI LAKLFPEYPVPTR
ShCCR2 SCCCCL6024F07.g 233 SYENAVAAAYVDVRDVARAHLVIERPDAARGRYLCIGAVLHRAELNAMIREFLPKYPVTAK
Os CCR 220 SYENAVAAAYVDVRDVARAHLVIERPDAARGRYLCIGAVLHRAELNAMIREFLPKYPVTAK

Sb CCR 299 CSDEVNPRKQPYKFSNQKLRDLGLEFRPYQSLSYDT/KNQPKSHLPVIGEQTTEADKEE
ShCCR1 SCCRZ2C01A04.g 296 CSDEVNPRKQPYKFSNQKLRDLGLEFRPYQSLSYDT/KNQPKSHLPVIGEQTTEADKEE
*Sr CCR 259 CSDEVNPRKQPYKFSNQSLVSPTSRTSS--QX-----
*Sba CCR 258 CSDEVNPRKQPYKFSNQSLV-----
*Ss CCR 257 CSDEVNPRKQPYKFSNQSLV-----
*So CCR 257 CSDEVNPRKQPYKFSNQSLV-----
ShCCR2 SCCCCL6024F07.g 293 CEDLGKEMAKPYRFSNQRLRDQGLEFTPRKSLYETVVCTCHKCHVFTKQKQRASL---
Os CCR 280 CEDLGKEMAKPYKFSNQRLKDLGLEFTPRKSLHDAVLCQDKSHLPITYPVPKRAYL--

Sb CCR 359 -ANAAAEVQGGIAIRA
ShCCR1 SCCRZ2C01A04.g 356 AAPAAAEVQGGIAIRA
*Sr CCR -----
*Sba CCR -----
*Ss CCR -----
*So CCR -----
ShCCR2 SCCCCL6024F07.g -----
Os CCR -----

```

**Figura S7A-F5H**

```

Sb      F5H      1  MAAVAKIAMEWLQDPLSCVFLVTLAVVLLQLRRRGKAPLPPGPKSLPIVGNMALMDQLTH
ShF5H1  SCJLRT1022E04.g 1  MAAVAKIAMEWLQDPLSCVFLVTLAVVLLQLRRRGKAPLPPGPKSLPIVGNMALMDQLTH
*Sba    F5H      1  -----SILMDQLTH
*So     F5H      1  -----EWRPREFDSLMDQLTH
*Ss     F5H      1  -----FTSDSMMDQLTH
*Sr     F5H      1  -----FTSDSLMDQLTH

Sb      F5H      61  RGLAALAAGDGVFSNRPATIAIAYLTYDRA
ShF5H1  SCJLRT1022E04.g 61  RGLAALAKEYGGLLHLRLGRLHAFVSTPEYAREVLQTQDGVFSNRPATIAIAYLTYDRA
*Sba    F5H      9  RGLAALKEYGGLLHLRLGRLHAFVSTPEYAREVLQTQDGVFSNRPATIAIAYLTYDRA
*So     F5H     17  RGLAALAKEYGGLLHLRLGRLHAFVSTPEYAREVLQTQDGVFSNRPATIAIAYLTYDRA
*Ss     F5H     13  RGLAALAKEYGGLLHLRLGRLHAFVSTPEYAREVLQTQDGVFSNRPATIAIAYLTYDRA
*Sr     F5H     13  RGLAALAKEYGGLLHLRLGRLHAFVSTPEYAREVLQTQDGVFSNRPATIAIAYLTYDRA

Sb      F5H     121  A--S
ShF5H1  SCJLRT1022E04.g 121  DMAFAHYGPFWRQMRKLCVMKLFSSRRRAETWVAVRDECAALVRGVAVSSGGGEKAVNLGE
*Sba    F5H     69  DMAFAHYGPFWRQMRKLCVMKLFSSRRRAETWVAVRDECAALVRGVAVSSGGGEKAVNLGE
*So     F5H     77  DMAFAHYGPFWRQMRKLCVMKLFSSRRRAETWVAVRDECAALVRGVAVSSGGGEKAVNLGE
*Ss     F5H     73  DMAFAHYGPFWRQMRKLCVMKLFSSRRRAETWVAVRDECAALVRGVAVSSGGGEKAVNLGE
*Sr     F5H     73  DMAFAHYGPFWRQMRKLCVMKLFSSRRRAETWVAVRDECAALVRGVAVSSGGGEKAVNLGE

Sb      F5H     179  LIFLTKNVTFRAAFGTRDGEDQEEFIAILQEFSKLFGAFNIGDFLPWLGWMDLQGINRR
ShF5H1  SCJLRT1022E04.g 181  LIFNLTKNVTFRAAFGTRDGEDQEEFIAILQEFSKLFGAFNIGDFLPWLGWMDLQGINRR
*Sba    F5H     129
*So     F5H     137
*Ss     F5H     133
*Sr     F5H     133

Sb      F5H     239  LRAARSALDRFIDKIIDEHVVRGKSPDDADADMDMLAFPAEAKPAAVNGGAAANGDDL
ShF5H1  SCJLRT1022E04.g 241  LRAARSALDRFIDKIIDEHVVRGKSPDDADADMDMLAFPAEAK-----
*Sba    F5H     189  LRAARSALDRFIDKIIDENHVRGRLQV---DHGELPTRNM-----
*So     F5H     197  LRAARSALDRFIDKIIDEQITSEFAAA-----CRSTINSSQVRVGC-----
*Ss     F5H     193  LRAARSALDRFIDK-----
*Sr     F5H     193  LRAARSALDRFIDKIDEQIT-----

Sb      F5H     299  QNTLRLTRDNIKAIIMDMVFGGTETVASAIEWAMAEMMHSPDDLRRVQQELADVGLDRN
ShF5H1  SCJLRT1022E04.g -----
*Sba    F5H -----
*So     F5H -----
*Ss     F5H -----
*Sr     F5H -----

Sb      F5H     359  VNESDLDKLPFLKCVIKETLRLHPPPIPLLLHETADDCVVGSGGGRRYSVPRGSRVMINVW
ShF5H1  SCJLRT1022E04.g -----
*Sba    F5H -----
*So     F5H -----
*Ss     F5H -----
*Sr     F5H -----

Sb      F5H     419  AIGRHRGSWKDADVFRPSRFTPDGDAAGLDFKGGCFEFLPFGSGRRSCPGTALGLYALEL
ShF5H1  SCJLRT1022E04.g -----
*Sba    F5H -----
*So     F5H -----
*Ss     F5H -----
*Sr     F5H -----

Sb      F5H     479  AVAQLAHGFNWSLPGMKPSELDMSDVFGLTAPRATRLYAVPTPRLNCPY
ShF5H1  SCJLRT1022E04.g -----
*Sba    F5H -----
*So     F5H -----
*Ss     F5H -----
*Sr     F5H -----

```

Figura S8A- COMT

```

Sb      COMT      1  MGSTAEDVAAVADEEACMYAMQLASSSILPMTLKNALELGLEVLQK--AGKALAAEEV
*So     COMT      1  -----LSILPMTLKNALELGLEVLQAEAPAGKALAPEEV
*Sba    COMT      1  -----HLSILPMTLKNALELGLEVLQAEAPAGKALAPEEV
*Ss     COMT      1  -----RIHLSILPMTLKNALELGLEVLQAEAPAGKALAPEEV
*Sr     COMT      1  ---CIQRVSSPIWSTC--RRPRIHLSILPMTLKNALELGLEVLQAEAPAGKALAPEEV
Ms      COMT      1  MGSTAEDVAAVADEEACMYAMQLASSSILPMTLKNALELGLEVLQAEAPAGKALAPEEV
ShCOMT1 SCJLRT1023B09.g 1  MGSTAEDVAAVADEEACMYAMQLASSSILPMTLKNALELGLEVLQAEAPAGKALAPEEV

Sb      COMT      59  VARLPVAPTNPDAADMVDRMLRLLASYDVVKCQMEDKDGKYERRYSAAPVGKWLTPNEDG
*So     COMT      36  VARLPVAPTNPDAADMVDRMLRLLASYDVVKCQMEDKDGKYERRYSAAPVGKWLTPNEDG
*Sba    COMT      37  VARLPVAPTNPDAADMVDRMLRLLASYDVVKCQMEDKDGKYERRYSAAPVGKWLTPNEDG
*Ss     COMT      39  VARLPVAPTNPDAADMVDRMLRLLASYDVVKCQMEDKDGKYERRYSAAPVGKWLTPNEDG
*Sr     COMT      56  VARLPVAPTNPDAADMVDRMLRLLASYDVVKCQMEDKDGKYERRYSAAPVGKWLTPNEDG
Ms      COMT      61  VARLPVAPTNPDAADMVDRMLRLLASYDVVKCQMEDKDGKYERRYSAAPVGKWLTPNEDG
ShCOMT1 SCJLRT1023B09.g 61  VARLPVAPTNPDAADMVDRMLRLLASYDVVKCQMEDKDGKYERRYSAAPVGKWLTPNEDG

Sb      COMT      119  VSMAALALMNQDKVLMESWYYLKDAVLDGGIPFNKAYGMTAFEYHGTDPRFNRVFNEMGK
*So     COMT      96  VSMAALALMNQDKVLMESWYYLKDAVLDGGIPFNKAYGMTAFEYHGTDPRFNRVFNEMGK
*Sba    COMT      97  VSMAALALMNQDKVLMESWYYLKDAVLDGGIPFNKAYGMTAFEYHGTDPRFNRVFNEMGK
*Ss     COMT      99  VSMAALALMNQDKVLMESWYYLKDAVLDGGIPFNKAYGMTAFEYHGTDPRFNRVFNEMGK
*Sr     COMT      116  VSMAALALMNQDKVLMESWYYLKDAVLDGGIPFNKAYGMTAFEYHGTDPRFNRVFNEMGK
Ms      COMT      121  VSMAALALMNQDKVLMESWYYLKDAVLDGGIPFNKAYGMTAFEYHGTDPRFNRVFNEMGK
ShCOMT1 SCJLRT1023B09.g121  VSMAALALMNQDKVLMESWYYLKDAVLDGGIPFNKAYGMTAFEYHGTDPRFNRVFNEMGK

Sb      COMT      179  NHSVITKKLLEFYTGFEESVSTLVDVGGGIGATLHAITSHHSHIRGINFDLPHVISEAP
*So     COMT      156  NHSVITKKLLEFYTGFE--GVSTLVDVGGGIGATLHAITSHHPQIKGINFDLPHVISEAP
*Sba    COMT      157  NHSVITKKLLEFYTGFE--GVSTLVDVGGGIGATLHAITSHHPQIKGINFDLPHVISEAP
*Ss     COMT      159  NHSVITKKLLEFYTGFE--GVSTLVDVGGGIGATLHAITSHHPQIKGINFDLPHVISEAP
*Sr     COMT      176  NHSVITKKLLEFYTGFE--GVSTLVDVGGGIGATLHAITSHHPQIKGINFDLPHVISEAP
Ms      COMT      181  NHSVITKKLLEFYTGFE--GVSTLVDVGGGIGATLHAITSHHPQIRGVNFDLPHVISEAP
ShCOMT1 SCJLRT1023B09.g181  NHSVITKK-----

Sb      COMT      239  PFPGVQHVGGDMFKSVPGDAILMKWILHDWSDAHCATLLKNCYDALPEKGGKVIVVECV
*So     COMT      215  PFPGVQHVGGDMFKSVPGDAILMKWILHDIIS-----
*Sba    COMT      216  PFPGVQHVGGDMFKSVPGDAILMKWILHD-----
*Ss     COMT      218  PFPGVQHVGGDMFKSVPGDAILMKWI-----
*Sr     COMT      235  PFPGVQHVGGDMFKSVPGDAILMKWILHDI-----
Ms      COMT      240  PFPVQHVGGDMFKSVPGDAILMKWILHDWSDAHCATLLKNCYDALPEN-GKVIVVECV
ShCOMT1 SCJLRT1023B09.g-----

Sb      COMT      299  LPVTTDAVPKAQGVFHVDMIMLAHNPGGRRERYERFRDLAKAAGFSGFKATYIYANAWAI
*So     COMT      -----
*Sba    COMT      -----
*Ss     COMT      -----
*Sr     COMT      -----
*Ms     COMT      299  LPVNTEAVPKAQGVFHVDMIMLAHNPGGRRERYERFRDLAKGAGFSGFKATYIYANAWAI
ShCOMT1 SCJLRT1023B09.g-----

Sb      COMT      359  EFIK
*So     COMT      ----
*Sba    COMT      ----
*Ss     COMT      ----
*Sr     COMT      ----
Ms      COMT      359  EFIK
ShCOMT1 SCJLRT1023B09.g-----

```





Pv CAD 2 180 L G S L L G H M G V K V A K A M G H H V T V I S S S S K K R A E A M H L G A D A Y L V S D A E A M A  
 \*So CAD A 139 P G L R G A I L G L G G V G H M G V K V A K A M G H H V T V I S S S S K K R A E A M H L G A D A Y L V S D A A A M A  
 \*Ss CAD A 139 P G L R G A I L G L G G V G H M G V K V A K A M G H H V T V I S S S S K K R A E A M H L G A D A Y L V S D A A A M A  
 \*Sr CAD A 137 P G L R G A I L G L G G V G H M G V K V A K A M G H H V T V I S S S S K K R A E A M H L G A D A Y L V S D A A A M A  
 \*Sba CAD A 139 P G L R G A I L G L G G V G H M G V K V A K A M G H H V T V I S S S S K K R A E A M H L G A D A Y L V S D A A A M A  
 ShCAD2 SCEPRZ1011A02.g 180 P G L R G A I L G L G G V G H M G V K V A K A M G H H V T V I S S S S K K R A E A M H L G A D A Y L V S D A A A M A  
 Zm CAD 180 P G L R G A I L G L G G V G H M G V K V A K A M G H H V T V I S S S S K K R A E A M H L G A D A Y L V S D A A A M G  
 ShCAD6 SCACHR1038E08.g 166 P G G S L G V V G L G G L G H V A V K F A K A F G L R V T V I S T S P A K R E A R E R L G A D H F V V S D Q K Q M Q  
 Sb CAD 182 A E A S L G V V G L G G L G H V A V K F A K A F G L R V T V I S T S P A K R E A R E R L G A D D F V V S D Q K Q M Q  
 \*Sr CAD C 141 K H I G V V G L G G L G H V A V K F A K A F G M R V T V I S T S P E K R E A M E T L G A D A F V V S D A N Q M K  
 ShCAD7 SCCCLB1001F10.g 183 P G K H I G V V G L G G L G H V A V K F A K A F G M R V T V I S T S P E K R E A M E T L G A D A F V V S D A N Q M K  
 Sb CAD 240 K H L G V V G L G G L G H V A V K F E K A F G M K V T V I S T S A S K R Q E A E N L G A D E F L I S R D P E E M K  
 \*Ss CAD B 141 A E K H L G V V G L G G L G H V A V K F E K A F G M K V T V I S T S A S K R Q E A E N L G A D E F L I S R D P E D M K  
 ShCAD8 SCEQLR1029E05.g 181 K H L G V V G L G G L G H V A V K F E K A F G M K V T V I S T S A S K R Q E A E N L G A D E F L I S R D P E Q M K  
 \*Sr CAD B 140 P G K H L G V V G L G G L G H V A V K F G K A F G M K V T V I S T S A S K R Q E A E N L G A D E F L I S R D P E Q M K  
 \*Sba CAD B 141 P G K H L G V V G L G G L G H V A V K F G K A F G M K V T V I S T S A S K R Q E A E N L G A D E F L I S R D P E Q M K  
 \*So CAD B 141 P G K H L G V V G L G G L G H V A V K F G K A F G M K V T V I S T S A S K R Q E A E N L G A D E F L I S R D P E Q M K

Pv CAD 2 240 A D S L D Y I I D T V P V H E P Y L A L L K L D G K H V L L G V I G E P L S F V A P M V M L G R K A T T G S F  
 \*So CAD A 199 A A A D S L D Y I I D T V P V H H P E P Y L A L L K L D G K H V L L G V I G E P L S F V S P M V M L G R K A T T G S F  
 \*Ss CAD A 199 A A A D S I D Y I I D T V P V H E P Y L A L L K L D G K H V L L G V I G E P L S F V S P M V M L G R K A T T G S F  
 \*Sr CAD A 197 A A A D S I D Y I I D T V P V H E P Y L A L L K L D G K H V L L G V I G E P L S F V S P M V M L G R K A T T G S F  
 \*Sba CAD A 199 E A A D S I D Y I I D T V P V H E P Y L A L L K L D G K H V L L G V I G E P L S F V S P M V M L G R K A T T G S F

ShCAD2 SCEPRZ1011A02.g 240 P A A D S I D Y I I D T V P V H E P Y L A L L K L D G K H V L L G V I G E P L S F V S P M V M L G R K A T T F  
 Zm CAD 226 A M A R S I D Y I I D T V S A R H S G P L E L L K V N G K L V L V A P D Q P V E L P S E P L I F G K T V S G S M  
 ShCAD6 SCACHR1038E08.g 242 A M A R S I D Y I I D T V S A K S G P L E L L K V N G K L V L V A P D Q P V E L P S E P L I F G K T V S G S M  
 Sb CAD 201 A V K G T M H G I N T A S A M S Y S Y L A L L K P H G K M I L G I P E R P L Q I S A S L V A G K T T A G S C  
 \*Sr CAD C 243 A V K G T M H G I N T A S A M S Y S Y L A L L K P H G K M I L G I P E R P L Q I S A S L V A G K T T A G S C  
 ShCAD7 SCCCLB1001F10.g 300 A A T G T M D B I I D T V S A W T P L A L L K P L G Q V I V G G S P S K P L E P A A I V P S K G V A G N S  
 Sb CAD 201 A A T G T M D B I I D T V S A W H P T P L A L L K P L G Q V I V G G S P S K P L E P V A F P S K G V A G N S  
 \*Ss CAD B 241 T G T M D B I I D T V S A W T L A L L K P L G Q V I V G G S P L E P V A F P S K G V A G N S  
 ShCAD8 SCEQLR1029E05.g 200 T G T M D B I I D T V S A W T L A L L K P L G Q V I V G G S P L E P A A F P S K G V A G N S  
 \*Sr CAD B 201 T G T M D B I I D T V S A W T L A L L K P L G Q V I V G G S P L E P A A F P S K G V A G N S  
 \*Sba CAD B 191 T G T M D B I I D T V S A W T L A L L K P L G Q V I V G G S P L E P A A F P S K G V A G N S

Pv CAD 2 300 G S I D T A E L L R F C V D K S L T S Q I E V K M G E R L E N D V R Y R F V D V A G S N I E E A T  
 \*So CAD A 259 G S I D T T E V L D F C V D K R L T S Q I E V K M G Y V N E A L D R L E N D V R Y R F V I D I E S L V -----  
 \*Ss CAD A 259 G S I D T A E V L D F C V D K S L T S Q I E V K M G Y V N E A L D R L E N D V R Y X F V I D I -----  
 \*Sr CAD A 257 G S I D T A E V L D F C V D K S L T S Q I E V K M G Y V N E A L D R L E N D V R Y R F V I D I E S L -----  
 \*Sba CAD A 259 G S I D T A E V L D F C V D K G L T S Q I E V K M G Y V N E A L D R L E N D V R Y R F V I D I E S L -----

ShCAD2 SCEPRZ1011A02.g 300 I I D T A E V L D F C V D K G L T S Q I E V K Y V N E A E R L E N D V R Y R F V D V A G S N V E A E A  
 Zm CAD 286 T G G M K T Q E M L L C Q H N I T C I E L V S T G I N E A L A R L A N D V R Y R F V I D I A G D S N S K L  
 ShCAD6 SCACHR1038E08.g 302 T G G M K T Q E M L L C Q H G I T C I E L V S T G I N A Q R L A S N D V R Y R F V I D I A G G D S S N S K  
 Sb CAD 261 M G S I R T Q G M M F A K H G T A I E L I G P E V N E A E R L A G V R Y R F V I D I -----  
 \*Sr CAD C 303 M G S I R T Q G M M F A K T A I E L I G P E V N E A E R L A G V R Y R F V I D I G N T L S A A S L  
 ShCAD7 SCCCLB1001F10.g 360 M G S V G C C Q A M L E F A K I G A E V E V Y V N T I E R L E N D V R Y R F V I D I A G S L G S A A  
 Sb CAD 261 M G S V G C C Q A M L E F A K I G A E V E V Y V N T I E R L E N D V R Y R F V I D I -----  
 \*Ss CAD B 301 M G S V G C C Q A M L E F A K I G A E V E V Y V N T I E R L E N D V R Y R F V I D V I G S L G S A A  
 ShCAD8 SCEQLR1029E05.g 260 M G S V G C C Q A M L E F A K I G A E V E V Y V N T I E R L E N D V R Y R F V I D I E S L V -----  
 \*Sr CAD B 261 M G S V G C C Q A M L E F A K I G A E V E V Y V N T I E R L E N D V R Y R F V I D I E S L V -----  
 \*Sba CAD B 251 M G S V G C C Q A M L E F A K I G A E V E V Y V N T I E R L E N D V R Y R F V I D I E S L V -----  
 \*So CAD B

Pv	CAD 2	360	GAPAN---
*So	CAD A		-----
*Ss	CAD A		-----
*Sr	CAD A		-----
*Sba	CAD A		-----
ShCAD2	SCEPRZ1011A02.g		-----
Zm	CAD	360	AAADAASN
ShCAD6	SCACHR1038E08.g		-----
Sb	CAD	362	L-----
*Sr	CAD C		-----
ShCAD7	SCCCLB1001F10.g	363	ASSPVPAL
Sb	CAD		-----
*Ss	CAD B		-----
ShCAD8	SCEQLR1029E05.g		-----
*Sr	CAD B		-----
*Sba	CAD B		-----
*So	CAD B		-----

## 5. REFERÊNCIAS BIBLIOGRÁFICAS

- Aden A, Foust T** (2009) Technoeconomic analysis of the dilute sulfuric acid and enzymatic hydrolysis process for the conversion of corn stover to ethanol. *Cellulose* **16**: 535–545
- Amaral LD, Gaspar M, Costa PMF, Aidar MPM, Buckeridge MS** (2007) Novo método enzimático rápido e sensível de extração e dosagem de amido em materiais vegetais. *Hoehnea* **34**: 425–431
- Anderson WF, Akin DE** (2008) Structural and chemical properties of grass lignocelluloses related to conversion for biofuels. *J Ind Microbiol Biotechnol* **35**: 355–366
- Anterola AM, Jeon J-H, Davin LB, Lewis NG** (2002) Transcriptional Control of Monolignol Biosynthesis in *Pinus taeda* factors affecting monolignol ratios and carbon allocation in phenylpropanoid metabolism. *J Biol Chem* **277**: 18272–18280
- Arai-Sanoh Y, Ida M, Zhao R, Yoshinaga S, Takai T, Ishimaru T, Maeda H, Nishitani K, Terashima Y, Gau M, et al** (2011) Genotypic variations in non-structural carbohydrate and cell-wall components of the stem in rice, sorghum, and sugar vane. *Biosci Biotechnol Biochem* **75**: 1104–12
- Austrália** (2008) The Biology of the *Saccharum* spp.
- Bacic A, Churms SC, Stephen AM, Cohen PB, Fincher GB** (1987) Fine structure of the arabinogalactan-protein from *Lolium multiflorum*. *Carbohydr Res* **162**: 85–93
- Bacic A, Harris PJ, Stone BA** (1988) Structure and function of plant cell walls. *Biochem plants* **14**: 297–371
- Balbo JM, Padovani Neto A** (1987) Excedentes de energia elétrica e sobra de bagaço para diferentes concepções e sistemas de conversão e utilização de energia aplicáveis a indústria sucro-alcooleira. *STAB, Açúcar, Álcool e Sub-produtos* **692**: 52–58
- Barrière Y, Riboulet C, Méchin V, Maltese S, Pichon M, Cardinal A, Lapierre C, Lubberstedt T, Martinant J-P** (2007) Genetics and genomics of lignification in grass cell walls based on maize as model species. *Genes Genomes Genomics* **1**: 133–156
- Bekker JPI** (2007) Genetic manipulation of the cell wall composition of sugarcane. Stellenbosch: University of Stellenbosch
- Berlyn GP, Miksche JP** (1976) Botanical Microtechnique and Cytochemistry. Iowa State University Press
- Bewley JD, Black M** (1994) Seeds; physiology of development and germination. New York: Plenum Press
- Blount JW, Korth KL, Masoud SA, Rasmussen S, Lamb C, Dixon RA** (2000) Altering expression of cinnamic acid 4-hydroxylase in transgenic plants provides evidence for a feedback loop at the entry point into the phenylpropanoid pathway. *Plant Physiol* **122**: 107–116
- Boerjan W, Ralph J, Baucher M** (2003) Lignin biosynthesis. *Annu Rev Plant Biol* **54**: 519–46

- Bolwell GP, Cramer CL, Lamb CJ, Schuch W, Dixon RA** (1986) L-Phenylalanine ammonia-lyase from *Phaseolus vulgaris*: Modulation of the levels of active enzyme by trans-cinnamic acid. *Planta* **169**: 97–107
- Bonawitz ND, Chapple C** (2010) The genetics of lignin biosynthesis: connecting genotype to phenotype. *Annu Rev Genet* **44**: 337–63
- Bottcher A, Cesarino I, Santos AB dos, Vicentini R, Mayer JLS, Vanholme R, Morreel K, Goeminne G, Moura JCMS, Nobile PM, et al** (2013) Lignification in sugarcane: biochemical characterization, gene discovery, and expression analysis in two genotypes contrasting for lignin content. *Plant Physiol* **163**: 1539–57
- Boudet A-M** (2000) Lignins and lignification: selected issues. *Plant Physiol Biochem* **38**: 81–96
- Bouveng HO** (1961) Phenylisocyanate derivatives of carbohydrates. 2. Location of o-acetyl groups in birch xylan. *Acta Chem Scand* **15**: 96
- Brown L, Torget R** (1996) NREL analytical procedure: LAP009 enzymatic saccharification of lignocellulosic biomass hydrolysis. Golden, CO Natl. Renew. Energy Lab.
- Buckeridge MS, Cavalari AA, Silva GB** (2008) Parede celular. *Fisiol Veg* 165–181
- Buckeridge MS, Santos WD dos, Sousa AP de** (2010) As rotas para o etanol celulósico no Brasil. Bioetanol da cana açúcar P&D para Prod e sustentabilidade 365–380
- Bull TA, Glasziou KT** (1963) The evolutionary significance of sugar accumulation in *Saccharum*. *Aust J Biol Sci* **16**: 737–742
- Burnquist WL, Sorrells ME, Tanksley S** (1992) Characterization of genetic variability in *Saccharum* germplasm by means of restriction fragment length polymorphism (RFLP) analysis. *Proc Int Soc Sugar Cane Technol* **21**: 355–365
- Butkutė, B., Lemežienė, N., Cesevičienė, J., Liatukas, Ž., & Dabkevičienė G** (2013) Carbohydrate and lignin partitioning in switchgrass *Panicum virgatum* biomass as a bioenergy feedstock. *Zemdirbyste-Agriculture* **100**: 251–260
- Buxton DR, Russell JR** (1988) Lignin constituents and cell-wall digestibility of grass and legume stems. *Crop Sci* **28**: 553–558
- Campbell MM, Sederoff RR** (1996) Variation in Lignin Content and Composition (Mechanisms of Control and Implications for the Genetic Improvement of Plants). *Plant Physiol* **110**: 3
- Carpita NC** (1989) Pectic polysaccharides of maize coleoptiles and proso millet cells in liquid culture. *Phytochemistry* **28**: 121–125
- Carpita NC** (1986) Incorporation of proline and aromatic amino acids into cell walls of maize coleoptiles. *Plant Physiol* **80**: 660–666
- Carpita NC** (1996) Structure and biogenesis of the cell walls of grasses. *Annu Rev Plant Biol* **47**: 445–476
- Carpita NC, Gibeaut DM** (1993) Structural models of primary cell walls in flowering plants:

consistency of molecular structure with the physical properties of the walls during growth. *Plant J* **3**: 1–30

**Carroll A, Somerville C** (2009) Cellulosic Biofuels. *Annu Rev Plant Biol* **60**: 165–182

**Carvalho-Netto O V, Bressiani JA, Soriano HL, Fiori CS, Santos JM, Barbosa GVS, Xavier MA, Landell MGA, Pereira GAG** (2014) The potential of the energy cane as the main biomass crop for the cellulosic industry. *Chem Biol Technol Agric* **1**: 1–8

**Chen C, Meyermans H, Burggraefe B, De Rycke RM, Inoue K, De Vleeschauwer V, Steenackers M, Van Montagu MC, Engler GJ, Boerjan WA** (2000) Cell-specific and conditional expression of caffeoyl-coenzyme A-3-O-methyltransferase in poplar. *Plant Physiol* **123**: 853–868

**Chen F, Dixon RA** (2007) Lignin modification improves fermentable sugar yields for biofuel production. *Nat Biotechnol* **25**: 759–761

**Chen F, Kota P, Blount JW, Dixon RA** (2001) Chemical syntheses of caffeoyl and 5-OH coniferyl aldehydes and alcohols and determination of lignin O-methyltransferase activities in dicot and monocot species. *Phytochemistry* **58**: 1035–1042

**Chen F, Yasuda S, Fukushima K** (1999) Evidence for a novel biosynthetic pathway that regulates the ratio of syringyl to guaiacyl residues in lignin in the differentiating xylem of *Magnolia kobus* DC. *Planta* **207**: 597–603

**Chen L, Auh C, Chen F, Cheng X, Aljoe H, Dixon RA, Wang Z** (2002) Lignin deposition and associated changes in anatomy, enzyme activity, gene expression, and ruminal degradability in stems of tall fescue at different developmental stages. *J Agric Food Chem* **50**: 5558–5565

**Cheng K, Sorek H, Zimmermann H, Wemmer DE, Pauly M** (2013) Solution-state 2D NMR spectroscopy of plant cell walls enabled by a dimethylsulfoxide-d 6/1-ethyl-3-methylimidazolium acetate solvent. *Anal Chem* **85**: 3213–3221

**Chong S-L, Virkki L, Maaheimo H, Juvonen M, Derba-Maceluch M, Koutaniemi S, Roach M, Sundberg B, Tuomainen P, Mellerowicz EJ, et al** (2014) O-Acetylation of glucuronoxylan in *Arabidopsis thaliana* wild type and its change in xylan biosynthesis mutants. *Glycobiology* cwu017

**D'Hont A, Ison D, Alix K, Roux C, Glaszmann JC** (1998) Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. *Genome* **41**: 221–225

**D'hont A, Lu YH, Feldmann P, Glaszmann JC** (1993) Cytoplasmic diversity in sugarcane revealed by heterologous probes. *Sugar cane* **1**: 12–15

**Daniels J, Roach BT** (1987) Taxonomy and evolution. *Sugarcane Improv. through breeding*. Elsevier, Amsterdam 7:

**Davin LB, Lewis NG** (2000) Dirigent proteins and dirigent sites explain the mystery of specificity of radical precursor coupling in lignan and lignin biosynthesis. *Plant Physiol* **123**: 453–462

**DuBois M, Gilles KA, Hamilton JK, Rebers PA, Smith F** (1956) Colorimetric Method for

- Determination of Sugars and Related Substances. *Anal Chem* **28**: 350–356
- Dutt NL, Narasimhan R** (1951) Starch in the genus *Saccharum* and interspecific and intergeneric hybrids. *Proc Sugarcane Res Work Indian Union* **1**: 4–10
- Eggleston G, Viator R, Grisham M** (2007) Glyphosate ripener effects on the processing quality of different sugarcane tissues. *Proc. Int. Soc. Sugar Cane Technol.* pp 1460–1467
- Ehlting J, Büttner D, Wang Q, Douglas CJ, Somssich IE, Kombrink E** (1999) Three 4-coumarate: coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. *Plant J* **19**: 9–20
- Farrell AE, Plevin RJ, Turner BT, Jones AD, O'Hare M, Kammen DM** (2006) Ethanol can contribute to energy and environmental goals. *Science* **311**: 506–8
- Ferrer J-L, Austin MB, Stewart C, Noel JP** (2008) Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiol Biochem* **46**: 356–370
- Fincher GB** (2009) Revolutionary times in our understanding of cell wall biosynthesis and remodeling in the grasses. *Plant Physiol* **149**: 27–37
- Fincher GB, Stone BA, Clarke AE** (1983) Arabinogalactan-proteins: structure, biosynthesis, and function. *Annu Rev Plant Physiol* **34**: 47–70
- Franke R, Humphreys JM, Hemm MR, Denault JW, Ruegger MO, Cusumano JC, Chapple C** (2002) The *Arabidopsis* REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism. *Plant J* **30**: 33–45
- Galletti GC, Muscarella CI, Bocchini P, Cinti F** (1996) Changes in the phenolic composition of maize stovers after blooming. *J Agric Food Chem* **44**: 3840–3844
- Gille S, de Souza A, Xiong G, Benz M, Cheng K, Schultink A, Reca I-B, Pauly M** (2011) O-acetylation of *Arabidopsis* hemicellulose xyloglucan requires AXY4 or AXY4L, proteins with a TBL and DUF231 domain. *Plant Cell* **23**: 4041–4053
- Glaszmann JC, Lu Y h, Lanaud C** (1990) Variation of nuclear ribosomal DNA in sugarcane. *J Genet Breed* **44**: 191–198
- Goldemberg J** (2007) Ethanol for a sustainable energy future. *Science* **315**: 808–10
- Grabber JH, Hatfield RD, Lu F, Ralph J** (2008) Coniferyl ferulate incorporation into lignin enhances the alkaline delignification and enzymatic degradation of cell walls. *Biomacromolecules* **9**: 2510–2516
- Grabber JH, Jung GA, Abrams SM, Howard DB** (1992) Digestion kinetics of parenchyma and sclerenchyma cell walls isolated from orchardgrass and switchgrass. *Crop Sci* **32**: 806–810
- Grabber JH, Quideau S, Ralph J** (1996) p-Coumaroylated syringyl units in maize lignin: Implications for  $\beta$ -ether cleavage by thioacidolysis. *Phytochemistry* **43**: 1189–1194
- Grivet L, Glaszmann JC, D'Hont A** (2006) Molecular evidence of sugarcane evolution and domestication. *Darwin's Harvest New approaches to Orig Evol Conserv Crop* Columbia Univ Press New York 49–66

- Gross GG** (1985) Biosynthesis and metabolism of phenolic acids and monolignols. Academic Press: Orlando, FL
- Gubler F, Ashford AE** (1985) Release of ferulic acid esters from barley aleurone. I. Time course of gibberellic-acid-induced release from isolated layers. *Funct Plant Biol* **12**: 297–305
- Guerra A, Filpponen I, Lucia LA, Argyropoulos DS** (2006) Comparative evaluation of three lignin isolation protocols for various wood species. *J Agric Food Chem* **54**: 9696–9705
- Guillaumie S, San-Clemente H, Deswarte C, Martinez Y, Lapierre C, Murigneux A, Barrière Y, Pichon M, Goffner D** (2007) MAIZEWALL. Database and developmental gene expression profiling of cell wall biosynthesis and assembly in maize. *Plant Physiol* **143**: 339–363
- Guo D, Chen F, Dixon RA** (2002) Monolignol biosynthesis in microsomal preparations from lignifying stems of alfalfa (*Medicago sativa* L.). *Phytochemistry* **61**: 657–667
- Hall TA** (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**: 95–98
- van Handel E** (1968) Direct microdetermination of sucrose. *Anal Biochem* **22**: 280–283
- Harding SA, Leshkevich J, Chiang VL, Tsai C-J** (2002) Differential substrate inhibition couples kinetically distinct 4-coumarate: coenzyme A ligases with spatially distinct metabolic roles in quaking aspen. *Plant Physiol* **128**: 428–438
- Harris PJ, Hartley RD** (1980) Phenolic constituents of the cell walls of monocotyledons. *Biochem Syst Ecol* **8**: 153–160
- Hatfield R, Fukushima RS** (2005) Can lignin be accurately measured? *Crop Sci* **45**: 832–839
- Hatfield RD, Marita JM, Frost K, Grabber J, Ralph J, Lu F, Kim H** (2009) Grass lignin acylation: p-coumaroyl transferase activity and cell wall characteristics of C3 and C4 grasses. *Planta* **229**: 1253–1267
- Hayashi T** (1989) Xyloglucans in the primary cell wall. *Annu Rev Plant Biol* **40**: 139–168
- He L, Terashima N** (1991) Formation and Structure of Lignin in Monocotyledons IV. Deposition Process and Structural Diversity of the Lignin in the CellWall of Sugarcane and Rice Plant Studied by Ultraviolet Microscopic Spectroscopy. *Holzforschung-International J Biol Chem Phys Technol Wood* **45**: 191–198
- Helle S, Cameron D, Lam J, White B, Duff S** (2003) Effect of inhibitory compounds found in biomass hydrolysates on growth and xylose fermentation by a genetically engineered strain of *S. cerevisiae*. *Enzyme Microb Technol* **33**: 786–792
- Hisano H, Nandakumar R, Wang Z-Y** (2009) Genetic modification of lignin biosynthesis for improved biofuel production. *Vitr Cell Dev Biol - Plant* **45**: 306–313
- Hoffmann L, Maury S, Martz F, Geoffroy P, Legrand M** (2003) Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in



phenylpropanoid metabolism. *J Biol Chem* **278**: 95–103

**Huang HJ, Ramaswamy S, Al-Dajani W, Tschirner U, Cairncross RA** (2009) Effect of biomass species and plant size on cellulosic ethanol: A comparative process and economic analysis. *Biomass and Bioenergy* **33**: 234–246

**Hughes J, McCully ME** (1975) The use of an optical brightener in the study of plant structure. *Stain Technol* **50**: 319–29

**Humphreys JM, Chapple C** (2002) Rewriting the lignin roadmap. *Curr Opin Plant Biol* **5**: 224–229

**Humphreys JM, Hemm MR, Chapple C** (1999) New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. *Proc Natl Acad Sci* **96**: 10045–10050

**Iiyama K, Lam TBT, Meikle PJ, Ng K, Rhodes DI, Stone BA** (1993) Cell wall biosynthesis and its regulation. *Forage cell wall Struct Dig* 621–683

**Iiyama K, Lam TBT, Stone BA** (1994) Covalent cross-links in the cell wall. *Plant Physiol* **104**: 315

**Iiyama K, Lam TBT, Stone BA** (1990) Phenolic acid bridges between polysaccharides and lignin in wheat internodes. *Phytochemistry* **29**: 733–737

**Imrie FKE, Tilbury RH** (1972) Polysaccharides in sugar cane and its products. *Sugar Technol Rev* **1**: 291–361

**Inman-Bamber G, Jackson P, Bonnett G, Morgan T** (2011) Have we reached peak CCS? *Int Sugar J* **113**: 798

**Ishii T** (1991) Isolation and characterization of a diferuloyl arabinoxylan hexasaccharide from bamboo shoot cell-walls. *Carbohydr Res* **219**: 15–22

**Jarvis MC** (1984) Structure and properties of pectin gels in plant cell walls. *Plant Cell Environ* **7**: 153–164

**Jin H, Cominelli E, Bailey P, Parr A, Mehrtens F, Jones J, Tonelli C, Weisshaar B, Martin C** (2000) Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in *Arabidopsis*. *EMBO J* **19**: 6150–6161

**Johansen DA** (1940) *Plant microtechnique*. McGraw-Hill Publishing Company, Ltd., London

**Jones L, Ennos AR, Turner SR** (2001) Cloning and characterization of irregular xylem4 (*irx4*): a severely lignin-deficient mutant of *Arabidopsis*. *Plant J* **26**: 205–216

**Jung HG, Casler MD** (2006) Maize Stem Tissues: Cell Wall Concentration and Composition during Development. *Crop Sci* **46**: 1793–1800

**Jung H-JG, Samac DA, Sarath G** (2012) Modifying crops to increase cell wall digestibility. *Plant Sci* **185**: 65–77

**Kishimoto T, Chiba W, Saito K, Fukushima K, Uraki Y, Ubukata M** (2009) Influence of syringyl to guaiacyl ratio on the structure of natural and synthetic lignins. *J Agric Food Chem* **58**: 895–901

- Kiyota E, Mazzafera P, Sawaya ACHF** (2012) Analysis of soluble lignin in sugarcane by ultrahigh performance liquid chromatography-tandem mass spectrometry with a do-it-yourself oligomer database. *Anal Chem* **84**: 7015–20
- Klein-Marcuschamer D, Oleskiewicz-Popiel P, Simmons BA, Blanch HW** (2010) Technoeconomic analysis of biofuels: A wiki-based platform for lignocellulosic biorefineries. *biomass and bioenergy* **34**: 1914–1921
- Klemm D, Philipp B, Heinze T, Heinze U, Wagenknecht W** (1998) *Comprehensive Cellulose Chemistry: Volume I: Fundamentals and analytical Methods*.
- Lam TBT, Iiyama K, Stone BA** (1992) Cinnamic acid bridges between cell wall polymers in wheat and phalaris internodes. *Phytochemistry* **31**: 1179–1183
- Lamb C** (1977) trans-Cinnamic acid as a mediator of the light-stimulated increase in hydroxycinnamoyl-CoA: Quinate hydroxycinnamoyl transferase. *FEBS Lett* **75**: 37–40
- Lan W, Lu F, Regner M, Zhu Y, Rencoret J, Ralph SA, Zakai UI, Morreel K, Boerjan W, Ralph J** (2015) Tricin, a flavonoid monomer in monocot lignification. *Plant Physiol* **167**: 1284–1295
- Laureano-Perez L, Teymouri F, Alizadeh H, Dale BE** (2005) Understanding factors that limit enzymatic hydrolysis of biomass. *Appl Biochem Biotechnol* **124**: 1081–1099
- Lavarack BP, Griffin GJ, Rodman D** (2002) The acid hydrolysis of sugarcane bagasse hemicellulose to produce xylose, arabinose, glucose and other products. *Biomass and Bioenergy* **23**: 367–380
- Lennartsson PR, Erlandsson P, Taherzadeh MJ** (2014) Integration of the first and second generation bioethanol processes and the importance of by-products. *Bioresour Technol* **165**: 3–8
- Li L, Cheng XF, Leshkevich J, Umezawa T, Harding SA, Chiang VL** (2001) The last step of syringyl monolignol biosynthesis in angiosperms is regulated by a novel gene encoding sinapyl alcohol dehydrogenase. *Plant Cell* **13**: 1567–1586
- Li L, Popko JL, Umezawa T, Chiang VL** (2000) 5-Hydroxyconiferyl aldehyde modulates enzymatic methylation for syringyl monolignol formation, a new view of monolignol biosynthesis in angiosperms. *J Biol Chem* **275**: 6537–6545
- Li X, Weng J-K, Chapple C** (2008) Improvement of biomass through lignin modification. *Plant J* **54**: 569–81
- Liepmann AH, Wightman R, Geshi N, Turner SR, Scheller HV** (2010) Arabidopsis - A powerful model system for plant cell wall research. *Plant J* **61**: 1107–1121
- Lindberg B, Rosell KG, Svensson S** (1973) Positions of the O-acetyl groups in birch xylan. *Sven. Papperstidn*
- Lindermayr C, Möllers B, Fliegmann J, Uhlmann A, Lottspeich F, Meimberg H, Ebel J** (2002) Divergent members of a soybean (*Glycine max* L.) 4-coumarate: coenzyme A ligase gene family. *Eur J Biochem* **269**: 1304–1315
- Lingle SE, Thomson JL** (2012) Sugarcane internode composition during crop development.

BioEnergy Res **5**: 168–178

- Lu F, Ralph J** (2002) Preliminary evidence for sinapyl acetate as a lignin monomer in kenaf. *Chem Commun* 90–91
- Lu F, Ralph J** (1999) Detection and determination of p-coumaroylated units in lignins. *J Agric Food Chem* **47**: 1988–1992
- Lynd LR, Wyman CE, Gerngross TU** (1999) Biocommodity engineering. *Biotechnol Prog* **15**: 777–793
- Magid J, Luxhøi J, Lyshede OB** (2004) Decomposition of plant residues at low temperatures separates turnover of nitrogen and energy rich tissue components in time. *Plant Soil* **258**: 351–365
- Marita JM, Ralph J, Hatfield RD, Chapple C** (1999) NMR characterization of lignins in *Arabidopsis* altered in the activity of ferulate 5-hydroxylase. *Proc Natl Acad Sci* **96**: 12328–12332
- Markwalder HU, Neukom H** (1976) Diferulic acid as a possible crosslink in hemicelluloses from wheat germ. *Phytochemistry* **15**: 836–837
- Marrone L** (2015) Metabolismo de Lignina em quatro genótipos de Cana-de-açúcar cultivados em campo. Universidade estadual de Campinas, São Paulo-Brasil
- Masarin F, Gurpilhares DB, Baffa DCF, Barbosa MHP, Carvalho W, Ferraz A, Milagres AMF** (2011) Chemical composition and enzymatic digestibility of sugarcane clones selected for varied lignin content. *Biotechnol Biofuels* **4**: 1
- Mäule C** (1901) Das verhalten verholzter membranen gegen kaliumpermanganat, eine holzreaktion neuer art. A. Zimmer's Verlag (Ernst Mohrmann)
- McCann MC, Roberts K** (1991) Architecture of the primary cell wall. Cytoskeleton basis plant growth form 109–129
- McCormick AJ, Cramer Md, Watt DA** (2006) Sink strength regulates photosynthesis in sugarcane. *New Phytol* **171**: 759–770
- Meijer AH, Cardoso M, Voskuilen JT, Waal A, Verpoorte R, Hoge JHC** (1993) Isolation and characterization of a cDNA clone from *Catharanthus roseus* encoding NADPH: cytochrome P-450 reductase, an enzyme essential for reactions catalysed by cytochrome P-450 mono-oxygenases in plants. *Plant J* **4**: 47–60
- Ming R, Moore PH, Wu KK, D'hont A, Glaszmann JC, Tew TL, Mirkov TE, da Silva J, Jifon J, Rai M, et al** (2006) Sugarcane Improvement through Breeding and Biotechnology. *Plant Breed. Rev.* pp 15–118
- Mizutani M, Ohta D, Sato R** (1997) Isolation of a cDNA and a genomic clone encoding cinnamate 4-hydroxylase from *Arabidopsis* and its expression manner in planta. *Plant Physiol* **113**: 755–763
- Mokochinski JB, Bataglion GA, Kiyota E, de Souza LM, Mazzafera P, Sawaya ACHF** (2015) A simple protocol to determine lignin S/G ratio in plants by UHPLC-MS. *Anal Bioanal Chem* **407**: 7221–7227

- Moore KJ, Jung H-JG** (2001) Lignin and fiber digestion. *J range Manag* 420–430
- Moore PH** (1987) Anatomy and morphology. *Sugarcane Improv through Breed* 85–142
- Morreel K, Dima O, Kim H, Lu F, Niculaes C, Vanholme R, Dauwe R, Goeminne G, Inzé D, Messens E, et al** (2010) Mass spectrometry-based sequencing of lignin oligomers. *Plant Physiol* **153**: 1464–1478
- Morrison TA, Kessler JR, Buxton DR** (1994) Maize internode elongation patterns. *Crop Sci* **34**: 1055–1060
- Moura JCMS, Bonine CAV, de Oliveira Fernandes Viana J, Dornelas MC, Mazzafera P** (2010) Abiotic and biotic stresses and changes in the lignin content and composition in plants. *J Integr Plant Biol* **52**: 360–76
- Naidu KM, Sreenivasan T V** (1987) Conservation of Sugarcane germplasm. *Copersugar Int. Sugarcane Breed. Work. Piracicaba-SP, Brazil*, pp 3–52
- Nelson N** (1944) A photometric adaptation of the somogyi method for the determination of glucose. *J Biol Chem* **153**: 375–380
- de Obeso M, Caparrós-Ruiz D, Vignols F, Puigdomènech P, Rigau J** (2003) Characterisation of maize peroxidases having differential patterns of mRNA accumulation in relation to lignifying tissues. *Gene* **309**: 23–33
- De Oliveira MED, Vaughan BE, Rykiel EJ** (2005) Ethanol as Fuel: Energy, Carbon Dioxide Balances, and Ecological Footprint. *Bioscience* **55**: 593
- Osakabe K, Tsao CC, Li L, Popko JL, Umezawa T, Carraway DT, Smeltzer RH, Joshi CP, Chiang VL** (1999) Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms. *Proc Natl Acad Sci* **96**: 8955–8960
- Pan Y-Q, Luo H-L, Li Y-R** (2009) Soluble acid invertase and sucrose phosphate synthase: Key enzymes in regulating sucrose accumulation in sugarcane stalk. *Sugar Tech* **11**: 28–33
- Papini-Terzi FS, Rocha FR, Vêncio RZN, Felix JM, Branco DS, Waclawovsky AJ, Del Bem L, Lembke CG, Costa MDL, Nishiyama MY, et al** (2009) Sugarcane genes associated with sucrose content. *BMC Genomics* **10**: 1
- Pauly M, Keegstra K** (2008) Cell-wall carbohydrates and their modification as a resource for biofuels. *Plant J* **54**: 559–568
- Pauly M, Keegstra K** (2010) Plant cell wall polymers as precursors for biofuels. *Curr Opin Plant Biol* **13**: 304–311
- Pawar PM-A, Koutaniemi S, Tenkanen M, Mellerowicz EJ, others** (2013) Acetylation of woody lignocellulose: significance and regulation. *Front Plant Sci* **4**: 118–126
- Pena CG, Mitrani RB, Correa JL, Cadenas GA, Munillas MH** (2000) Bagasse, in *Handbook of sugarcane derivatives*. ICIDCA. Bugallo, S R Havana City, Cuba, pp 29–40
- Pinçon G, Chabannes M, Lapierre C, Pollet B, Ruel K, Joseleau J-P, Boudet AM,**

- Legrand M** (2001) Simultaneous down-regulation of caffeic/5-hydroxy ferulic acid-O-methyltransferase I and cinnamoyl-coenzyme A reductase in the progeny from a cross between tobacco lines homozygous for each transgene. Consequences for plant development and lignin synthesis. *Plant Physiol* **126**: 145–155
- Porto BN, Magalhães PC, Campos NA, Alves JD, Magalhães MM** (2011) Otimização de protocolos de extração de rna em diferentes tecidos de milho. *Rev Bras Milho e Sorgo* **9**: 189–200
- Price S** (1963) Cytogenetics of modern sugar canes. *Econ Bot* **17**: 97–106
- Puigdomenech PC, Perez P, Murigneux A, Martinant JP, Tixier MH, Rigau J, Civardi L, Maes T** (2001) Identifying genes associated with a QTL corn digestibility locus. *Pat. WO 0155395-A 1* 3:
- Raes J, Rohde A, Christensen JH, de Peer Y, Boerjan W** (2003) Genome-wide characterization of the lignification toolbox in Arabidopsis. *Plant Physiol* **133**: 1051–1071
- Ralph J, Akiyama T, Kim H, Lu F, Schatz PF, Marita JM, Ralph SA, Reddy MSS, Chen F, Dixon RA** (2006) Effects of coumarate 3-hydroxylase down-regulation on lignin structure. *J Biol Chem* **281**: 8843–53
- Ralph J, Hatfield RD, Quideau S, Helm RF, Grabber JH, Jung H-JG** (1994a) Pathway of p-coumaric acid incorporation into maize lignin as revealed by NMR. *J Am Chem Soc* **116**: 9448–9456
- Ralph J, Helm RF, Quideau S, Hatfield RD** (1992) Lignin-feruloyl ester cross-links in grasses. Part 1. Incorporation of feruloyl esters into coniferyl alcohol dehydrogenation polymers. *J Chem Soc Perkin Trans 1* **21**: 2961–2969
- Ralph J, Lundquist K, Brunow G, Lu F, Kim H, Schatz PF, Marita JM, Hatfield RD, Ralph SA, Christensen JH, et al** (2004) Lignins: natural polymers from oxidative coupling of 4-hydroxyphenyl-propanoids. *Phytochem Rev* **3**: 29–60
- Ralph J, MacKay JJ, Hatfield RD, O'Malley DM, Whetten RW, Sederoff RR** (1997) Abnormal lignin in a loblolly pine mutant. *Science* (80- ) **277**: 235–239
- Ralph J, Quideau S, Grabber JH, Hatfield RD** (1994b) Identification and synthesis of new ferulic acid dehydrodimers present in grass cell walls. *J Chem Soc Perkin Trans 1* **23**: 3485–3498
- Ramakrishnan C V** (1958) Sucrose and starch synthesis in sugar cane plant. *Experientia* **14**: 91–92
- Ramos LP** (2003) The chemistry involved in the steam treatment of lignocellulosic materials. *Quim Nova* **26**: 863–871
- Rancour DM, Marita JM, Hatfield RD** (2012) Cell wall composition throughout development for the model grass *Brachypodium distachyon*. *Front Plant Sci* **3**: 266
- Rass-Hansen J, Falsig H, Jørgensen B, Christensen CH** (2007) Bioethanol: fuel or feedstock? *J Chem Technol Biotechnol* **82**: 329–333

- Ren R** (2012) global status report 2012. Renewable energy policy network for the 21st century. Paris: REN 21 Secretariat
- Rencoret J, Marques G, Gutierrez A, Ibarra D, Li J, Gellerstedt G, Santos JI, Jimenez-Barbero J, Martínez AT, del Río JC** (2008) Structural characterization of milled wood lignins from different eucalypt species. *Holzforschung* **62**: 514–526
- Rencoret J, Marques G, Gutiérrez A, Nieto L, Jiménez-Barbero J, Martínez ÁT, José C** (2009) Isolation and structural characterization of the milled-wood lignin from *Paulownia fortunei* wood. *Ind Crops Prod* **30**: 137–143
- del Río JC, Prinsen P, Rencoret J, Nieto L, Jiménez-Barbero J, Ralph J, Martínez AT, Gutiérrez A** (2012) Structural characterization of the lignin in the cortex and pith of elephant grass (*Pennisetum purpureum*) stems. *J Agric Food Chem* **60**: 3619–3634
- Roach BT** (1978) Utilization of *Saccharum spontaneum* in sugarcane breeding. *Proc Int Soc Suagr Cane Technol* **16**: 43–58
- Roach BT, Daniels J** (1987) The *Saccharum* Complex and the genus *Saccharum*. *Copersugar Int. Sugarcane Breed. Work.* (Copersucar eds), Brazil, pp 1–33
- Rodrigues J, Meier D, Faix O, Pereira H** (1999) Determination of tree to tree variation in syringyl/guaiacyl ratio of *Eucalyptus globulus* wood lignin by analytical pyrolysis. *J Anal Appl Pyrolysis* **48**: 121–128
- Rudall PJ, Caddick LR** (1994) Investigation of the presence of phenolic compounds in monocotyledonous cell walls, using UV fluorescence microscopy. *Ann Bot* **74**: 483–491
- Ruth MF, Thomas SR** (2003) Effect of Corn Stover Composition on Ethanol Process Economics. 25th Symp. Biotechnol. Fuels Chem. NREL Publications, Breckenridge, Colorado, p 22
- Saha BC** (2003) Hemicellulose bioconversion. *J Ind Microbiol Biotechnol* **30**: 279–291
- Santos AB dos, Bottcher A, Kiyota E, Mayer JLS, Vicentini R, Brito M dos S, Creste S, Landell MGA, Mazzafera P** (2015) Water Stress Alters Lignin Content and Related Gene Expression in Two Sugarcane Genotypes. *J Agric Food Chem* **63**: 4708–4720
- dos Santos AB, Bottcher A, Vicentini R, Mayer JLS, Kiyota E, Landell MAG, Creste S, Mazzafera P** (2015) Lignin biosynthesis in sugarcane is affected by low temperature. *Environ Exp Bot* **120**: 31–42
- Santos RB, Hart P, Jameel H, Chang H** (2013) Wood based lignin reactions important to the biorefinery and pulp and paper industries. *BioResources* **8**: 1456–1477
- Saska M, Ozer E** (1995) Aqueous extraction of sugarcane bagasse hemicellulose and production of xylose syrup. *Biotechnol Bioeng* **45**: 517–523
- Sato P.** (2012) Regulação do acúmulo de sacarose em cana de açúcar e análise funcional de uma proteína quinase relacionada com o conteúdo de sacarose. Universidade de São Paulo, São Paulo-Brasil.
- Sato S, Kato T, Kakegawa K, Ishii T, Liu Y-G, Awano T, Takabe K, Nishiyama Y, Kuga S, Sato S, et al** (2001) Role of the putative membrane-bound endo-1, 4- $\beta$ -glucanase

- KORRIGAN in cell elongation and cellulose synthesis in *Arabidopsis thaliana*. *Plant Cell Physiol* **42**: 251–263
- Scalbert A, Monties B, Lallemand J-Y, Guittet E, Rolando C** (1985) Ether linkage between phenolic acids and lignin fractions from wheat straw. *Phytochemistry* **24**: 1359–1362
- Scheller HV, Ulvskov P** (2010a) Hemicelluloses. *Annu Rev Plant Biol* **61**: 263–289
- Scheller HV, Ulvskov P** (2010b) Hemicelluloses. *Plant Biol* **61**: 263
- Schoch G, Goepfert S, Morant M, Hehn A, Meyer D, Ullmann P, Werck-Reichhart D** (2001) CYP98A3 from *Arabidopsis thaliana* is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. *J Biol Chem* **276**: 36566–36574
- Scobbie L, Russell W, Provan GJ, Chesson A** (1993) The newly extended maize internode: a model for the study of secondary cell wall formation and consequences for digestibility. *J Sci Food Agric* **61**: 217–225
- Shen H, Fu C, Xiao X, Ray T, Tang Y, Wang Z, Chen F** (2009) Developmental control of lignification in stems of lowland switchgrass variety Alamo and the effects on saccharification efficiency. *BioEnergy Res* **2**: 233–245
- Shen H, Mazarei M, Hisano H, Escamilla-Trevino L, Fu C, Pu Y, Rudis MR, Tang Y, Xiao X, Jackson L, et al** (2013) A genomics approach to deciphering lignin biosynthesis in switchgrass. *Plant Cell* **25**: 4342–4361
- Shibuya N, Nakane R** (1984) Pectic polysaccharides of rice endosperm cell walls. *Phytochemistry* **23**: 1425–1429
- da Silva AM** (2005) Caracterização da parede celular de *Saccharum officinarum* L.(cana-de-açúcar) e *Brachiaria decumbens* Stapf (braquiária). Universidade estadual de Campinas. pp 102
- Silva GB, Ionashiro M, Carrara TB, Crivellari AC, Tiné MAS, Prado J, Carpita NC, Buckeridge MS** (2011) Cell wall polysaccharides from fern leaves: Evidence for a mannan-rich Type III cell wall in *Adiantum raddianum*. *Phytochemistry* **72**: 2352–2360
- Simmons BA, Loque D, Blanch HW** (2008) Next-generation biomass feedstocks for biofuel production. *Genome Biol* **9**: 242
- Slewinski TL** (2012) Non-structural carbohydrate partitioning in grass stems: a target to increase yield stability, stress tolerance, and biofuel production. *J Exp Bot* **63**: 24
- Sobral BW, Braga DP, Lahood ES, Keim P** (1994) Phylogenetic analysis of chloroplast restriction enzyme site mutations in the *Saccharinae* Griseb. subtribe of the *Andropogoneae* Dumort. tribe. *Theor Appl Genet* **87**: 843–53
- Socol CR, Vandenberghe LP de S, Medeiros ABP, Karp SG, Buckeridge M, Ramos LP, Pitarelo AP, Ferreira-Leitão V, Gottschalk LMF, Ferrara MA, et al** (2010) Bioethanol from lignocelluloses: Status and perspectives in Brazil. *Bioresour Technol* **101**: 4820–4825
- Somerville C, Youngs H, Taylor C, Davis SC, Long SP** (2010) Feedstocks for

lignocellulosic biofuels. *Science* **329**: 790–2

- Somogyi M** (1937) A reagent for the copper-iodometric determination of very small amounts of sugar. *J Biol Chem* **117**: 771–776
- Sousa CR** (2011) Caracterização da mobilização dos polissacarídeos da parede celular em palhada de cana-de-açúcar submetida às condições de campo. Universidade de São Paulo, São Paulo, 2011
- de Souza AP, Grandis A, Leite DCC, Buckeridge MS** (2014) Sugarcane as a bioenergy source: history, performance, and perspectives for second-generation bioethanol. *BioEnergy Res* **7**: 24–35
- de Souza AP, Leite DCC, Pattathil S, Hahn MG, Buckeridge MS** (2013) Composition and structure of sugarcane cell wall polysaccharides: implications for second-generation bioethanol production. *Bioenergy Res* **6**: 564–579
- Spiridon I, Popa VI** (2005) Hemicelluloses: structure and properties. *Polysaccharides Struct. Divers. Funct. versatility* **2**:
- Sun R, Tomkinson J, Zhu W, Wang SQ** (2000) Delignification of maize stems by peroxymonosulfuric acid, peroxyformic acid, peracetic acid, and hydrogen peroxide. 1. Physicochemical and structural characterization of the solubilized lignins. *J Agric Food Chem* **48**: 1253–1262
- Sun S-N, Li M-F, Yuan T-Q, Xu F, Sun R-C** (2012) Sequential extractions and structural characterization of lignin with ethanol and alkali from bamboo (*Neosinocalamus affinis*). *Ind Crops Prod* **37**: 51–60
- Sun X-F, Sun R-C, Zhao L, Sun J-X** (2004) Acetylation of sugarcane bagasse hemicelluloses under mild reaction conditions by using NBS as a catalyst. *J Appl Polym Sci* **92**: 53–61
- Suprasanna P, Patade VY, Desai NS, Devarumath RM, Kavar PG, Pagariya MC, Ganapathi A, Manickavasagam M, Babu KH** (2011) Biotechnological developments in sugarcane improvement: an overview. *Sugar Tech* **13**: 322–335
- Swain T, Hillis WE** (1959) The phenolic constituents of *Prunus domestica*. I.—The quantitative analysis of phenolic constituents. *J Sci Food Agric* **10**: 63–68
- TAPPI** (1985) TAPPI Useful Method UM 250: Acid-soluble lignin in wood and plants. TAPPI Useful Methods
- Tew T, Cobill R** (2008) Genetic Improvement of Sugarcane (*Saccharum* spp.) as an Energy Crop. *In* W Vermerris, ed, *Genet. Improv. Bioenergy Crop*. SE - 9. Springer New York, pp 273–294
- Thomas JR, Darvill AG, Albersheim P** (1989) Rhamnogalacturonan I, a pectic polysaccharide that is a component of monocot cell-walls. *Carbohydr Res* **185**: 279–305
- Trabucco GM, Matos DA, Lee SJ, Saathoff AJ, Priest HD, Mockler TC, Sarath G, Hazen SP** (2013) Functional characterization of cinnamyl alcohol dehydrogenase and caffeic acid O-methyltransferase in *Brachypodium distachyon*. *BMC Biotechnol* **13**: 1



- Tsuji Y, Chen F, Yasuda S, Fukushima K** (2005) Unexpected behavior of coniferin in lignin biosynthesis of *Ginkgo biloba* L. *Planta* **222**: 58–69
- Tsuji Y, Chen F, Yasuda S, Fukushima K** (2004) The behavior of deuterium-labeled monolignol and monolignol glucosides in lignin biosynthesis in angiosperms. *J Agric Food Chem* **52**: 131–134
- Tsuji Y, Fukushima K** (2004) Behavior of monolignol glucosides in angiosperms. *J Agric Food Chem* **52**: 7651–7659
- Turner LB, Mueller-Harvey I, McAllan AB** (1993) Light-induced isomerization and dimerization of cinnamic acid derivatives in cell walls. *Phytochemistry* **33**: 791–796
- Ulbrich B, Zenk MH** (1980) Partial purification and properties of p-hydroxycinnamoyl-CoA: shikimate-p-hydroxycinnamoyl transferase from higher plants. *Phytochemistry* **19**: 1625–1629
- Vijayan Nair N, Nair S, Sreenivasan T V., Mohan M** (1999) Analysis of genetic diversity and phylogeny in *Saccharum* and related genera using RAPD markers. *Genet Resour Crop Evol* **46**: 73–79
- Waclawovsky AJ, Sato PM, Lembke CG, Moore PH, Souza GM** (2010) Sugarcane for bioenergy production: an assessment of yield and regulation of sucrose content. *Plant Biotechnol J* **8**: 263–76
- Wang J, Nayak S, Koch K, Ming R** (2013) Carbon partitioning in sugarcane (*Saccharum* species). *Front Plant Sci* **4**: 201
- Wilson JR, Hatfield RD** (1997) Structural and chemical changes of cell wall types during stem development: consequences for fibre degradation by rumen microflora. *Aust J Agric Res* **48**: 165–180
- Wu X, Wu J, Luo Y, Bragg J, Anderson O, Vogel J, Gu YQ** (2013) Phylogenetic, molecular, and biochemical characterization of caffeic acid o-methyltransferase gene family in *Brachypodium distachyon*. *Int J Plant Genomics* **2013**: 12
- Xiong G, Cheng K, Pauly M** (2013) Xylan O-acetylation impacts xylem development and enzymatic recalcitrance as indicated by the *Arabidopsis* mutant *tbl29*. *Mol Plant* **6**: 1373–1375
- Xu F, Sun R-C, Zhai M-Z, Sun J-X, Jiang J-X, Zhao G-J** (2008) Comparative study of three lignin fractions isolated from mild ball-milled *Tamarix austromongolica* and *Caragana sepium*. *J Appl Polym Sci* **108**: 1158–1168
- Xu Z, Zhang D, Hu J, Zhou X, Ye X, Reichel KL, Stewart NR, Syrenne RD, Yang X, Gao P, et al** (2009) Comparative genome analysis of lignin biosynthesis gene families across the plant kingdom. *BMC Bioinformatics* **10**: 1
- Yamauchi K, Yasuda S, Fukushima K** (2002) Evidence for the biosynthetic pathway from sinapic acid to syringyl lignin using labeled sinapic acid with stable isotope at both methoxy groups in *Robinia pseudoacacia* and *Nerium indicum*. *J Agric Food Chem* **50**: 3222–3227
- Yasuda S, Fukushima K, Kakehi A** (2001) Formation and chemical structures of acid-

soluble lignin I: sulfuric acid treatment time and acid-soluble lignin content of hardwood. *J wood Sci* **47**: 69–72

**Ye Z-H, Kneusel RE, Matern U, Varner JE** (1994) An alternative methylation pathway in lignin biosynthesis in *Zinnia*. *Plant Cell* **6**: 1427–1439

**Yuan JS, Tiller KH, Al-Ahmad H, Stewart NR, Stewart CN** (2008) Plants to power: bioenergy to fuel the future. *Trends Plant Sci* **13**: 421–9

**Zhong R, Morrison WH, Negrel J, Ye Z-H** (1998) Dual methylation pathways in lignin biosynthesis. *Plant Cell* **10**: 2033–2045

**Zhou M, Kimbeng CA, Eggleston G, Veremis JC, Gravois KA** (2007) Prospects of breeding for low starch content in sugarcane. *Proc. Int. Soc. Sugar Cane Technol.* pp 724–729

**Zhou MM, Kimbeng CA, Eggleston G, Viator RP, Hale AL, Gravois KA** (2008) Issues of starch in sugarcane processing and prospects of breeding for low starch content in sugarcane. *Sugar Cane Int* **26**: 3

**Zubieta C, Kota P, Ferrer J-L, Dixon RA, Noel JP** (2002) Structural basis for the modulation of lignin monomer methylation by caffeic acid/5-hydroxyferulic acid 3/5-O-methyltransferase. *Plant Cell* **14**: 1265–1277

II

III

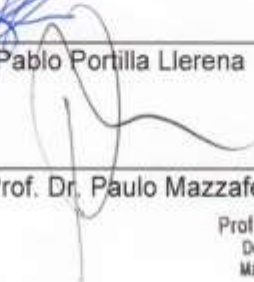
COORDENADORIA DE PÓS-GRADUAÇÃO  
INSTITUTO DE BIOLOGIA  
Universidade Estadual de Campinas  
Caixa Postal 6109, 13083-970, Campinas, SP, Brasil  
Fone (19) 3521-6378, email: cpgeb@unicamp.br



## DECLARAÇÃO

Em observância ao §5º do Artigo 1º da Informação CCPG-UNICAMP/001/15, referente a Bioética e Biossegurança, declaro que o conteúdo de minha Dissertação de Mestrado, intitulada "*Lignina em espécies de Saccharum spp.*", desenvolvida no Programa de Pós-Graduação em Biologia Vegetal do Instituto de Biologia da Unicamp, não versa sobre pesquisa envolvendo seres humanos, animais ou temas afetos a Biossegurança.

Assinatura:   
Nome do(a) aluno(a): Juan Pablo Portilla Llerena

Assinatura:   
Nome do(a) orientador(a): Prof. Dr. Paulo Mazzafera


Prof. Dr. Paulo Mazzafera  
Depto. Biologia Vegetal/IB  
Matricula Unicamp 19729-7

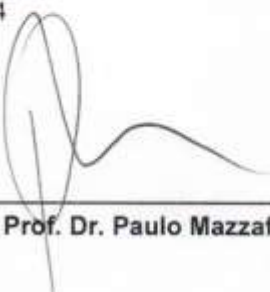
Data: 05 de Maio do 2016

**Profa. Dra. Rachel Meneguello**  
Presidente  
Comissão Central de Pós-Graduação  
**Declaração**

As cópias de artigos de minha autoria ou de minha co-autoria, já publicados ou submetidos para publicação em revistas científicas ou anais de congressos sujeitos a arbitragem, que constam da minha Dissertação/Tese de Mestrado/Doutorado, intitulada **Lignina em espécies de *Saccharum spp.***, não infringem os dispositivos da Lei n.º 9.610/98, nem o direito autoral de qualquer editora.

Campinas, 05 de Maio do 2016

Assinatura :   
Nome do(a) autor(a): **Juan Pablo Portilla Llerena**  
RG n.º Passport : 6579894

Assinatura :   
Nome do(a) orientador(a): **Prof. Dr. Paulo Mazzafera**  
RG n.º 8336840-1

Prof. Dr. Paulo Mazzafera  
Depto. Biologia Vegetal/B  
Matricula Unicamp 19729-7